International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2015): 6.391

Novel Approach to Virtual Machine Migration In Cloud Computing Environment – A Survey

Priyanka H[1], Dr. Mary Cherian[2]

[1]Research Scholar, CSE R&D Centre, Dr. Ambedkar Institute of Technology, (Affiliated to VTU), Bengaluru

[2]Associate Professor, CSE Department, Dr. Ambedkar Institute of Technology, Bengaluru, India

Abstract: ***Cloud computing is a powerful technology to perform large scale and complex computing. It abolishes the need to maintain expensive computing hardware, dedicated space, and software. Virtual Machine (VM) is an operating system or application that is installed on the software which has virtual dedicated hardware. Virtual machine migration helps in performing various management operations in data centers. These include load balancing, resource hotspot reduction, server coupling, and hardware/software up gradation. The movement of a virtual machine from one server (source) to another server (destination) is called VM migration. The migration process consumes a certain amount of CPU and network resources. The performance of the migration depends on (i) the migration time--the time between the start of the migration process and the end of the VM at the source, and (ii) the downtime-the time for which the virtual machine is paused. The cost metrics are CPU & network bandwidth utilization. In this paper, the evolution and need of VM Migration in cloud computing environment along with various approaches to select destination host are discussed***

Keywords: Virtual Machine, Virtual Machine Migration, Cloud Computing

1. Introduction

Cloud Computing is computing over internet is a best method of accessing and managing user data without any requirement of infrastructure at the user's location. Cloud user can access their data remotely from any location at a less cost and it guarantees the management and other security related issues. It has more advantages compared to its disadvantages. The major advantage of using cloud computing is that the resources are in sharing mode and can be accessed simultaneously by multiple users by using the concept of virtualization.

A. Virtualization

Virtualization is a most important feature of cloud computing to deals with number of devices and users over the network. It creates virtual environment on a single physical computer by abstracting the hardware details, which allows us to use multiple instances of operating systems (known as guest operating system) to handle number of processes simultaneously and separately by each guest OS (virtual machine). Virtualization also allows VM migration from one to another machine. Virtualization methodology divides the computer resources into different virtual machines by partitioning the hardware and software. Thus, it reduces overall cost in space, power and other infrastructure.

Hypervisor or VM manager is a core module or program which handles more than one virtual operating systems on a single host. It is a responsibility of hypervisor to provide resources and a processor to each virtual operating system on the same host. Hypervisor allocates required resources to each VM's. Two types of hypervisor are listed below.

Type 1, which is known as a bare-metal hypervisor and runs directly on top of hardware. It is also known as a hardware virtualization engine.

Type 2, which operates as an application on top of an existing operating system. It is also known as a software virtualization.

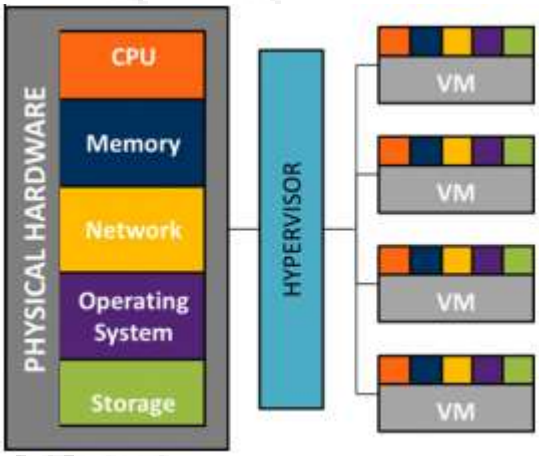

Figure 1: Hypervisor Layer

Virtualization technology allows the infrastructure provider to break up a physical machine into many smaller units of different sizes and offer them applications with different needs. These smaller units are well isolated from each other and execute their own copy of the operating system. They are also known as virtual machines (VMs). Each application is hosted inside a virtual machine to guarantee performance and fault isolation. Service Level Objectives (SLO) such as the 95th percentile response time or throughput can be specified for an application. Depending on the SLO, certain amount of physical resources such as CPU, memory, disk and network bandwidth is reserved for each application (virtual machine). One of the features of virtualization is live migration where a VM can be moved from one physical machine to another even as it continues to execute computing power by aggregating resources and offering a single system view. Depending on the SLO, certain amount of physical resources such as CPU, memory, disk and

network bandwidth is reserved for each application (virtual machine).

2. Necessity of VM Migration

The following are the main criteria for VM migration.

a) **Server coupling.** Servers are often under-utilized due to dynamic nature of the workload hosted on virtual machines. As a result, computing resources are idle and create inefficiency of resource usage in a data center. There are also implications on power consumption driven by the fact that even an idle computer consumes a certain amount of static power while doing no useful work. Server consolidation is an approach to efficiently pack virtual machines into a fewer number of servers to increase resource utilization and reduce power consumption. All virtual machines in the under-utilized server are moved to other servers using live migration technique.

b) **Load balancing.** This refers to distributing virtual machines across servers to balance the resource utilization across them. It results in maximum throughput, minimum response time, and avoidance of bottleneck at any single resource. As demand for workloads varies over time, load balancing approach periodically moves virtual machines to rebalance the load.

c) **Reducing resource hotspot.** Dynamic resource provisioning techniques allocate resources dynamically to a virtual machine based on changing load levels. However, if a server does not have enough free resources to satisfy a virtual machine's increased resource requirement, a resource hotspot results. A hotspot is resolved by migrating a virtual machine out of an overloaded physical machine. A hotspot is usually resolved, either by migrating the overloaded virtual machine to another physical machine that has adequate free resources, or by migrating any other virtual machines(s) to free up resources for the overloaded virtual machine.

d) **Mitigating interference between VMs**. Resources such as CPU, memory and network bandwidth can be partitioned and allocated to each virtual machine. Other resources such as CPU cache, disk and memory bandwidth cannot be partitioned and hence are shared amongst them. As a result, VM executions interfere with each other when they access these shared resources which degrades performance of application running on virtual machine. A solution to mitigating the effect of VM interference is to move interfering Vms apart and host them on separate Pms.

e) **Performing software & hardware upgrades**. Due to the large number of servers in data centers, the probability that some server will fail is fairly high. Further, software's or operating system executing on each server should be updated/patched periodically to fix security related issues. During these two scenarios, it is necessary to either shutdown the server (in the case of hardware upgrades) or reboot the server (in the case of software upgrades). To avoid a downtime for applications running on virtual machines, it is necessary to move virtual machines from such server to other server.

3. Literature Review

Although there are many approaches are introduced to enhance the existing live migration technology, they mainly focus on power-saving and load balancing scenarios.

In paper [1] VM bandwidth allocation is done in a fairway. Here the fairness is done with considering two main objectives:

1) Guaranteed bandwidth thereby maintaining SLA.
2) Sharing the residual bandwidth.

An algorithm called Falloc is introduced in this paper. The algorithm allocates a particular portion of the bandwidth on congested links to VMs based on the bandwidth requirements of these applications. The following important objectives are to be taken into consideration for developing the algorithm.

1) Bandwidth guarantee: It helps the VM to achieve high performance.
2) Weight assignment: The weight assigned should be different for different applications.

The idea of migrating OS instances for the better running of data center applications is introduced in [2]. Hence achieves a clear separation between the hardware and software and facilitates fault tolerant and seamless operation. This paper considers migrating OS instances in order to achieve the seamless communication. The commonly occurring interferences during multiple tasks execution is presented with several examples in [3]. This paper designs a prediction mechanism for a 4-dimensional resource model.

In [4] a system called sandpiper is introduced. This algorithm is used to find out hotspots and it determines a new mapping of physical to virtual resources and initiating the necessary VM Migration

4. Virtual Machine Migration Techniques

1) Non-Live Migration Technique

The Virtual Machine at the Source is stopped. All the memory pages are copied to destination host and the VM starts executing at the destination machine.

2) Live Migration

Live migration techniques are classified into two types. They are post-copy and pre-copy techniques.

a) Post-Copy Live Migration

In this approach VM stops execution on source machine and its execution state (CPU, register and Memory pages needed to activate VM on destination) is moved to destination and VM starts its execution on destination even the whole pages have not been copied. Page faults can occur if the required memory page of the VM is not present on the destination host, then the corresponding page is moved from source to destination. Figure 2 shows the working methodology of post-copy migration. The disadvantage of this method is the occurrence of page faults at the destination host.

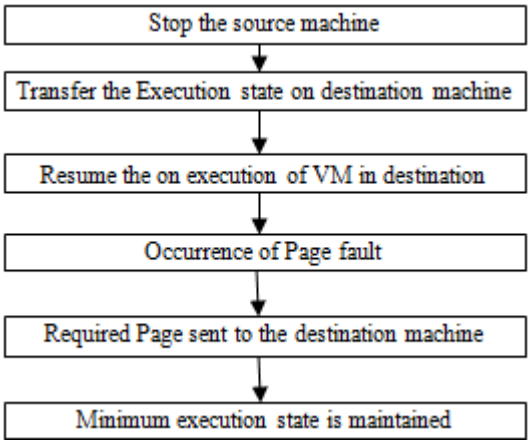

Figure 2: Flowchart of Post-copy migration

b) Pre-Copy Live Migration

Pre-copy is a live migration technique used by hypervisors like Xen, KVM and VMware. There are two phases in this method.

1) **Warm-up phase** -The memory pages are transferred from source machine to destination machine. The VM at the source is not stopped until required number of memory pages are copied to destination machine. If the pages are modified then it should be updated at the destination machine. The dirty page condition will occur if the rate of copied page to the target host is less than the rate of sent modified page.
2) **Stop and Copy phase** – After the completion of warm-up phase, the VM is stopped at the source machine and it starts executing at the destination machine. The disadvantage of this approach is that modified pages has to resend iteratively leading to inefficiency, it consumes more network bandwidth.

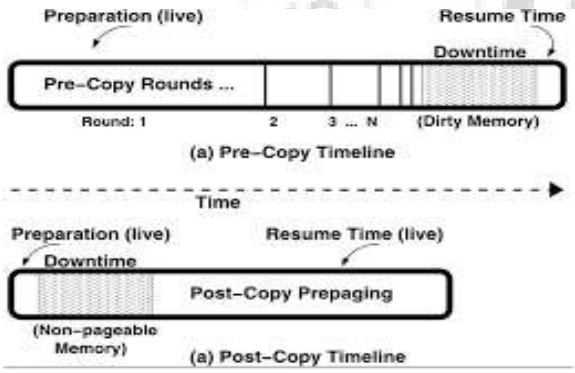

Figure 3: Pre-copy and Post-copy Migration

5. Techniques of Modified Pre-Copy Migration

1) Improved Pre-Copy Approach

It is the enhanced variant of by Pre-Copy approach by Ma et al[8]. It encourages the record to monitor habitually alter memory pages. To-send, To-skip, To-settle is the three bitmaps utilized by this approach. The pages changed in past emphasis can be spoken to by To-send and the present cycle altered pages can be signified by To-skip. To-settle holds the quantity of settled pages which are sent in the past round.

The as often as possible changed past round pages can be put away by To-send-last additional bitmap. The superfluous exchange of much of the time refreshed pages can be limited in this approach so the quantity of emphases is lessened. Every now and again adjusted pages can be resolved in the view of just To-send and To-skip bitmaps. It is not helpful for option altered pages in the memory amid relocation.

2) Matrix Bitmap Algorithm

Framework bit delineate defeats the disadvantage of improved pre-copy based approach. Numerous bitmaps are utilized to decide the memory page exchange between the source and goal. The quantity of bitmap gathered at every emphasis can be detached by a variable called MAP-LEN. Grimy bitmap gathered MAP-LEN times.

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 1 & & 1 \\ & \vdots & & \cdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \begin{matrix} P_1 \\ P_2 \\ \vdots \\ P_N \end{matrix}$$

In every emphasis unmodified pages can be controlled by zero and changed page dictated by one. Page weight is ascertained in view of the bitmap incentive to contrast and the edge esteem. The page with lesser weight than the edge can be sent in this emphasis. The limit is computed by

$$2^{MAP-LEN-1}$$

Also, if the weight is lesser than the limit then it can be sent to the goal. Outline is utilized to decide the edge esteem so; the achievement of this calculation relies on upon this variable. The limit esteem must be ideal for creating preferable results over the above pre-copy technique.

3) Two-Phase Strategy

The approach given by Ma et al. [8] is one phase strategy that sends the pages depends on previous iteration flag not supported the present iteration. However , the 2 phase strategy identifies high dirty pages by giving them Second probability (SC) to the page within the initial section To-send and To-skip are verified. If each the values are adequate to one and zero severally then it passes to second probability. If the page is clean during this phase then it's migrated to the destination. The duplication of frequently updated pages is avoided by 2 phase strategy. The steps to follow this technique are: at the start migration method activates supported the strategy of second probability. If the page is clean for each the iterations transmit the dirty page to focus on machine. The tip condition checks the amount of iterations or compares the count of duplicated pages to modify to at least one section strategy. SC strategy is often performed in twenty eight iterations although the optimum range of iterations for SC strategy lies in between one to twenty nine. Once this SC strategy switched to at least one section and performs stop and duplicate. If dirty page count is tiny than fifty five it's switched to the one phase as a result of the dirty number is small. Third condition verifies the dirty page count with memory pages of two and half size of VM. If condition satisfies then it switches to OP strategy

4) Pre-Copy using Memory Compression

Memory pressure gives quick and stable VM movement, proposed by jin H et al [6]. Trademark Based Compression (CBC) calculation is utilized to pack the VM memory pages before sending to the goal. This is one of the strategies to decrease the down time and aggregate relocation time. A decent compression calculation gives a higher compression

proportion and low overhead. The measure of packed information is signified by the compression proportion. The time taken to pack the memory pages by CPU is called overhead. High compression proportion and less overhead are the key parameters for memory compression calculation 6. The information is ordered into three gatherings in light of this calculation.

1) The memory page contains both zero and non-zero bytes. The esteem and the non-zero bytes are exchange to the objective when the zero byte check is more noteworthy than the zero byte limits.
2) The pressure of page will take less time when the page contains comparable words more prominent than the limit of comparable words.
3) Generally utilize the calculation to pack the page with high pressure proportion. Compression and decompression of a memory is overhead in this approach. Unused pages introduce in the memory are additionally compacted and sent to destination. The relocation time and overhead increment because of the superfluous transmission of indistinguishable pages.

The following table gives the summary of Virtual Machine migration techniques and their drawbacks.

Techniques/Algorithms	Future work and/or gaps in existing Technologies
Improved Time series based live migration technique	It works very well with structured and unstructured data. It is not applied to the alternative modified pages.
Bitmap algorithm with help of matrix	When the pages having weight more than the threshold value will not be forwarded to the Destination host. Strength of the algorithm is only based on the threshold value.
Fast live migration	Pages of the user processes are given page frame number (PFN) and transmitted from source to destination using storage area network I/o penalty will be the drawback of this approach which affects the QOS
Two Phase Pre-copy strategy	Improves the algorithm to meet the constraint of a given downtime for a certain class of applications

6. Conclusion

Live migration technique is used for balancing load in data centers without interrupting other running virtual machines. This paper illustrates the various pre-copy based live migration techniques which contributed in minimizing the migration time, downtime and increasing the performance. Each proposed method will contain its own drawback and introduces an overhead. It is tough to judge the different proposed methods as each of them is implemented in various architecture and platforms. It is also hard to standardize the performance of the migration

References

[1] J. Guo, F. Liu, H. Tang, Y. Lian, H. Jin, and J. C. Lui, "Falloc: Fair network bandwidth allocate on in IaaS datacenters via a bargaining game approach," in Proc. 21st IEEE Int. Conf. Netw Protocols (ICNP), 2013, pp. 7–10.
[2] C. Clark, K. Fraser, S. Hand, J.G. Hansen, E. Jul, C. Limpach, I. Pratt,and A. Warfield, "Live migration of virtual machines," in Proc. 2ndSymp. Netw. Syst. Design Implementation (NSDI'05), May 2005,pp. 273-286
[3] Christopher D. Wickens, "Multiple resources and performance prediction"University of Illinois at Urbana-Champaign, Institute of Aviation Willard Airport, AviationHuman Factors Division, Aviation Research Laboratory, 1 Airport Road, Savoy, IL 61874,USA, pp. 449-455.
[4] T. Wood, P. Shenoy, A. Venkataramani, and M. Yousif, "Black-box and gray-box strategies for virtual machine migration," in Proc. 2NdSymp. Netw. Syst. Design Implementation (NSDI), Apr. 2007, p. 17.
[5] Sagana C, Geetha M, Suganthe R. Performance enhance-ment in live migration for cloud computing environments. International Conference on Information Communication and Embedded Systems (ICICES); 2013
[6] Jin H, Deng L, Wu S, Shi X, Pan X. Live virtual machine migration with adaptive, memory compression. IEEE International Conference on Cluster Computing and Workshops (CLUSTER'09); 2009
[7] Rajakumari SB, Nalini C. An efficient cost model for data storage with horizontal layout in the cloud. Indian Journal of Science and Technology. 2014; 7(3S):45–6.
[8] Ma F, Liu F, Liu Z. Live virtual machine migration based on improved pre-copy approach. IEEE International Conference on Software Engineering and Service Sciences (ICSESS); 2010.