Computer Engineering and Intelligent Systems
ISSN 2222-1719 (Paper) ISSN 2222-2863 (Online)
Vol.6, No.6, 2015

Memory Management and Reuse Mechanism for Virtual Machine in Cloud Computing to Minimize Energy Consumption : A Review Paper

Sandeep Kaur
Department of Computer Science and Engineering,
Punjab Institute of Technology (PTU Main Campus), PTU University, Punjab India
Sandykaur374@gmail.com

Abstract
Cloud computing is an emerging computing technology for large data centers that maintains computational resources through the internet, rather than on local computers. VM migration provides the capability to balance the load, system maintenance and fault tolerance, etc. However, existing migration techniques, used to migrate virtual machines keeping memory images of VMs in host and skipping transfer of unchanged memory pages to reduce the amount of transfer data during migration, if number of migrations increases, number of memory images stored on host are also increased, this causes memory starvation. In this paper, a propose technique that reduces the size of memory image stored on source host before migration. When a VM migrates to other host, memory images of VM is kept in the source host after removing unwanted data according to the Probability factor. When the VM migrates back to the original host later, the kept memory image will be "reused", i.e. data which are identical to the kept data will not be transferred and comparative to existing system the size of memory image is small. To validate this approach, evaluate the results using different threshold levels and probability factor of change in data. Proposed system required less memory to store the memory image and allow more VMs to be hosted. Specifically, proposed work is used to improve resource efficiency throughout by reducing the size of memory image that is stored on source host.
Keywords: Cloud computing, Migration, Virtualization, Virtual Machine, Physical Machine, Live Virtual Machine Migration.

1. Introduction

Over the recent decades, propels in computing technology have revolutionized our lifestyle. Cloud computing engages IT holdings as an organization; suppliers are making a bestowed pool of configurable processing assets, which customers can alertly procurement and release on requests payable reason.

Dr. Rajkumar Buyya says [5] "A Cloud is a type of parallel and distributed system consisting of a collection of inter-connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resource(s) based on service-level agreements established through negotiation between the service provider and consumers." [3] .In cloud computing various cloud consumers demand variety of services as per their dynamically changing needs. So it is the job of cloud computing to avail all the demanded services to the cloud consumers. But due to the availability of finite resources it is very difficult for cloud providers to provide all the demanded services in time. From the cloud providers' perspective cloud resources must be allocated in a fair manner. So, it's a vital issue to meet cloud consumers' QoS requirements and satisfaction. The ultimate goal of efficient resource utilization in cloud computing is to maximize the profit for cloud providers and to minimize the cost for cloud consumers

Cloud computing is not new to Information Technology. It has been for the most part considered as an unlimited scale dispersed registering perfect model dictated by economy of scale. The peculiarities that perceive distributed computing from different models are administration driven, asset pooling, and information facilitating in outsourcing storage. Pooling asset makes the equipment execution be utilized more effective and gives monetary profits to clients to lessen the capital expense and extra use. The greatest profit is that engineers

no more oblige the vast capital expenses in hardware to deploy the inventive plans for new Internet administrations administration and subsequently slicing the human cost to work it [1] [23].

In this period where processing is offered as a utility, Datacenter is a key component encouraging endeavor applications. Each application is associated with a Service Level Agreement (SLA) organizing the application execution levels guaranteed to client and punishment installment structures in the occasion of the datacenter overseer's bafflement to respect SLA

The rest of the paper is organized as follows: Section 2 discusses Virtualization. Section 3 gives summary of virtual machine migration. Section 4 discusses various live virtual machine migration techniques proposed by researchers'. Section 5 presents conclusion.

2. Virtualization

Virtualization is a framework or methodology of dividing the resources of a computer into multiple execution environments, by applying one or more concepts or (VM) and technologies such as time-sharing, hardware and software partitioning, partial or complete machine simulation or emulation, quality of service, and many others [11]. The approach of virtualization which empowers to computing resources of a solitary physical machine (PM) among various virtual machines (VM) ensuring execution detachment, made ready for compelling and productive resource utilization and management. In a virtualized data center, every application segment (server) of an enterprise customer application is exemplified in a virtual machine (VM) and a solitary physical machine (PM) has different VMs. Virtualization gives an approach to change resources assigned to VMs dynamically (VM resizing) and moving VMs starting with one physical machine then onto the next. Furnished with an intelligent situation of virtual machines on physical machines, this permits us to take after the workload progressions of applications consequently empowering successful use of resources. Virtualization could be attained at the different type of levels. The Sorts of virtualizations are Server Virtualization, Storage Virtualization, Operating framework Virtualization and Network Virtualization, System Virtualization is the point at which a solitary physical host runs various VMs on it. This VM has it applications that run on its OS (guest OS). For the client, a VM carries on much the same as an autonomous physical machine.

3. VM Migration

Migration is one of the most important features provided by modern VM technologies. It allows system administrators to move an OS instance to another physical node without interrupting any hosted services on the migrating OS. It is an extremely powerful cluster administration tool and serves as a basis for many modern administration frameworks which aim to provide efficient online system maintenance, load balancing, reconfiguration and proactive fault tolerance in clusters and data-centres.

VM migration consumes resources both on the machine from which it is being migrated and the machine to which it is being migrated. As a result, performance of the VM in migration and VMs residing on both source and destination machines will be affected. The technique to determine and carry out this dynamic (re)allocation of VMs in a virtualized datacentre is termed as dynamic server consolidation. In addition to monitoring the resource utilization of VMs, any such technique needs to be wary of the effect of VM migrations too, to figure out the new VM-to-PM allocation map from time to time. There is a lot of work done on this area trying to devise an effective dynamic server consolidation technique.

Most of existing techniques try to migrate VMs from lightly loaded and heavily PMs to moderately used PMs While keeping number migrations to a minimum (to curb migrations cost). This allow to power of least utilized machines while decreasing load on highly utilized machines to prevent possible SLA violation. Each of them formulates migration cost in their own way (e.g. number of migrations, migration efficiently). Though the original aim of server consolidation was to minimize server sprawl - i.e. minimize the number of PMs being used, this should not be exercised naively. The main motive of data centre operator is to maximize its returns. Minimizing the number of active PMs obviously contributes to reduction in operational cost of data center. But this minimization carried out by migrating virtual machines is not free. Migration exhibits its own cost in terms of deteriorating application performance and reduction in amount of resources available while VM migrations are being done. If the deteriorated application performance meant a violation in service level agreement (SLA), penalty needs to be paid to client. Even then, migration may be a safe bet provided the aftermath reaps more revenue and this may not be possible unless previous migration activity is done.

4. VM migration Techniques

Initially, pure stop-and-copy approach was used for VM migration. This involves halting the original VM and copying all pages to the destination, after then starting the new VM. This has the advantages in terms of simplicity but the service downtime is proportional to the amount of physical memory which is allocated to the VM. This can be lead to an unacceptable outage if the VM is running a live service.

4.1 Pre-Copy Memory Migration

4.1.1 Warm-up Phase

In pre-copy memory migration, the Hypervisor typically copies all the memory pages from source to destination while the VM is still running on the source. If some memory pages change (become 'dirty') during this process, they will continue re-copied until the rate of re-copied pages is not less than page dirtying rate.

4.1.2 Stop-and-Copy Phase

After the warm-up phase, the VM will be stopped on the original host and the remaining dirty pages will be copied to the destination, and then the VM will be resumed on the destination host. The time gap between stopping the VM on the original host and resuming it on destination is called "down-time", and ranges from a few milliseconds to seconds according to the size of memory and applications running on the VM. There are

some techniques to reduce the migration down-time, such as by using probability density function of memory change.

4.2 Post-copy memory migration

Post-copy VM migration is initiated by suspending the VM at the source, by suspending the VM a minimal subset of the execution state of the VM (CPU registers and non-pageable memory) is transferred to the target. The VM is then resumed at the target, even though most of the memory state of the VM still resides at the source. At the target, when the VM tries to access pages that have not yet been transferred, it generates page-faults. These two faults are trapped at the target and redirected towards the source over the network. Such faults are referred to as network faults. The source host responds to the network-fault by sending the faulted page. Since each page fault of the running VM is redirected towards the source, this technique can degrade performance of applications running inside the VM. However, pure demand-paging accompanied with techniques such as pre-paging can reduce this impact by a great extent.

5. Literature Review

Sujesha Sudevalayam and Purushottam Kulkarni [7] argued that network affinity-awareness is required in resource provisioning for virtual machines. Authors have quantified their work of benchmarking of link network usage for both Xen and KVM virtualization technologies. Authors have also focused on building affinity-aware models that can predict expected CPU resource requirements based on its location relative to its communicating set of virtual machines – upon colocation and dispersion of virtual machines.

Pablo Graubner, Matthias Schmidt and Bernd Freisleben [8] presented a novel approach to virtual machine consolidation based on energy efficient storage migration and live VM migration. The Authors tried to save energy through virtual machine consolidation in IaaS cloud computing platform. The Authors had implemented the same approach using Eucalyptus which is an open source clone of the Amazon Elastic Compute Cloud (Amazon EC2).

Tiago C. Ferreto, Marco A. S. Netto, Rodrigo N. Calheiros and Cesar A.F. De Rose [10] a new approach named Dynamic Consolidation with various Migration Control. Authors have to discussed that current techniques like static consolidation and dynamic consolidation doesn't consider the steady usage of virtual machines. For which problems may arise as like migration cost and penalty to the physical server and the Authors claimed that for steady usage migration can be avoided but for variable usage migration can be performed. Authors had used Linear Programming Formulation and Heuristics approach for the same and the authors had evaluated their proposed approach with TU-Berlin Workload and Google Workload.

Mayank Mishra, Anwesha Das, Purushottam Kulkarni and Anirudha Sahoo [12] discussed that live virtual machine migration plays a vital role in dynamic resource management of cloud computing platform. Authors are mainly focused on efficient resource utilization in non peak periods to minimize wastages of resources. In the favour to achieve goals like hotspot mitigation, load balancing and the server consolidation, authors discussed mainly three components – when have to migrate, which VM have to migrate and where have to migrate – and approaches followed by different heuristics to apply various migration techniques. The Authors discussed the virtual machine migration over LAN and WAN with their challenges.

Haikum Liu, Hai Jin, Xiaotei Liao, Chen Yu and Cheng-Zhong Xu [13] had designed, implemented and evaluated a novel approach that minimises virtual machine migration downtime and network traffic. Authors had adopted check pointing/recovery and trace/reply technologies for the same and implemented a transparent virtual machine checkpoint with copy-on-write (COW) mechanism. Authors claimed that their proposed method can be used in both LAN and WAN. The experimental results showed that a novel approach gives better performance.

Kejiang Ye, Xiaohong Jiang, Dawei Huang, Jianhai Chen and Bei Wang [14] proposed resource reservation based live migration framework of multiple virtual machines. The proposed machine in the framework holds four virtual machines: Migration Controller, Migration Decision Maker, Resource Monitor and Resource Reservation Controller. Authors target on improving the migration efficiency through live migration of virtual machines and proposed three optimization methods: optimization in the parallel migration of multiple virtual machines, source machine and workload-aware migration strategy. To upgrade the migration efficiency authors had considered parameters like total migration time, downtime and workload performance on high. Authors declared that resource reservation strategy is required at source machine and target machine.

Febio Checconi, Tomasso Cucunotta and Manuel Stein [15] addressed real time issues in Live Virtual Machine Migration. Authors had presented a technique for live migration of real time applications. The main factor that had been considered by authors is down time of virtual machines due to live migration. The Authors have announced probabilistic model for migration process to find out new set of migration policies by building a sound mathematical theory. Deeper evaluation of the fully implemented proposed technique is still has to be done by the authors.

Soramichi Akiyama∗, Takahiro Hirofuchi†, Ryousei Takano† and Shinichi Honiden [9] discuss the

placement of Virtual Machines (VMs) on physical hosts are dynamically optimized in response to resource utilization of the physical hosts. However, the existing live migration techniques are used to move VMs from one host to another, and need to involve large data transfer and also prevents dynamic consolidation systems from optimizing VM placements efficiently. The proposed technique called "memory reusing "that reduces the amount of transferred memory of live migration. When a VM migrates between hosts, the memory image of the VM is kept on the original host (from where VM migrates). When the VM migration go back to the original host later, the kept memory mage will be "reused",i.e.memory pages which are identical to the kept pages will not be transferred. Implemented a system named MiyakoDori that uses memory reusing in live migrations.

Table 1. Summary of VM Migration Techniques

Year	Author	Techniques/Algorithms	Tools and/or workload	Future work and/or gaps in existing technologies
2013	Sujesha Sudevalayam and Purushottam Kulkarni [7]	Affinity aware modelling of CPU usage with communicating VMs	Xen and KVM	Heterogeneity of PMs are not considered. Benchmarking of 100 Mbps link network usage is done (1 Gbps – not done
2013	Pablo Graubner, Matthias Schmidt and Freisleben [8] Bernd	VM consolidation through energy efficient storage migration and live VM migration	Eucalyptus	Resource management at higher layer and overhead VM live migrations
Year	**Author**	**Techniques/Algorithms**	**Tools and/or workload used**	**Future work and/or gaps in existing technologies**
2011	Tiago C. Ferreto, Marco A. S. Netto, Rodrigo N. Calheiros and Cesar A.F. De Rose [10]	Dynamic server consolidation with migration control.	TU-Berlin and Google Workload	Can easily be implemented on VMware and Citrix Tools
2012	Mayank Mishra, Anwesha Das, Purushottam Kulkarni and Anirudha Sahoo [12]	Live virtual machine migration	Not Mentioned	Only load on the virtual machine for migration is considered. Consumer requirements and priority of job is not considered
2012	Soramichi Akiyama⋆, Takahiro Hirofuchi†, Ryousei Takano† and Shinichi Honiden [9]	Memory reusing mechanism for Dynamic VM consolidation	QEMU/KVM	Reduces the amount of transferred memory and total migration time of a live migration and thus reduces the energy consumption of a dynamic VM consolidation system.
2011	Haikum Liu, Hai Jin, Xiaotei Liao, Chen Yu and Cheng-Zhong Xu [13]	Virtual machine Checkpoint with copy-on write mechanism	On LAN and WAN	Multiprocessor virtual machine migration and design a hybrid scheme that can apply heuristics to choose alternative algorithm between pre-copy and proposed method
2011	Kejiang Ye, Xiaohong Jiang, Dawei Huang, Jianhai Chen and Bei Wang [14]	Live Migration of Virtual Machines	Xen and VMware	Intelligent live migration machine can be future work
2009	Febio Checconi, Tomasso Cucunotta and Manuel Stein [15]	Real Time Issues in Live Virtual Machine Migration (new set of migration policies)	KVM	Deeper evaluation of the fully implemented proposed technique is still has to be done by the authors

7. Conclusion

Research in the field of Virtualization has been in existence for more than a decade now as not only the technology is best suited for scaling demands for the IT infrastructure, It reduces the maintenance thereafter and

overall expenditure on infrastructure developments. With the developments in Virtual Machines design technology for both software and hardware, it now became easy to deploy new systems on existing ones without any hefty additional costs, maintain them, test new developments and provide best of the class service in potential application areas of Virtual Machines with migration process. The current work focuses on efficient memory management for source host and will reduce the memory requirement while migration. The main idea behind this concept is, whenever a VM is ready for migration from one host to another, its current state is stored on source host for future use to reduce the data transfer when it come back from the destination host after processing[9]. Miyakodori provides a new way in VM Migration to reduce the cost of data transfer, but it increases the memory consumption of the source host. The image stored on source host block the memory upto the completion of migration process and if the large number of virtual machines initialized from a single host, multiple images stored on source host which occupy a large chunk of memory, which may arise a situation of memory starvation if the memory images occupy most of the memory. At this state, host is not able to initialize a new virtual machine and no virtual machine will allowed on this host, which is coming from another host with a resource request. In our proposed system, a new technique will be implemented to overcome this starvation problem. According to this technique the size of memory image stored on the source host will be reduced. To reduce the size we applying to conditions:

The results will be better by using proposed technique. Further we implement this technique and compare the results with the existing technique.

Acknowledgment
The paper has been written with the kind assistance, guidance and active support of my department who have helped me in this work. I would like to thank all the individuals whose encouragement and support has made the completion of this work possible.

References

[1] Inderjit Singh Dhanoa, 2Dr. Sawtantar Singh Khurmi, "Energy-Efficient Virtual Machine Live Migration in Cloud Data Centers", ISSN: 0976-8491 (Online) | ISSN : 2229-4333 (Print), IJCST Vol. 5, SPL - 1, Jan - March 2014

[2] Shweta Rajput, Trupti manik, "Live-VM Migration Policies, Attacks & Security – A Survey", Volume 4, Issue 2, February 2014 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering

[3] Esha Barlaskar, Y. Jayanta Singh, "Dynamic Consolidation of Virtual Machines with Multi-Agent System", International Journal of Computer Applications (0975 – 8887), Volume 94 – No 9, May 2014

[4] Jiaxing Song, Weidong Liu, Feiran Yin, and Chao Gao, "TSMC: A Novel Approach for Live Virtual Machine Migration", Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China, Received 23 January 2014; Accepted 6 May 2014; Published 20 May 2014.

[5] Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopala, James Broberg, Ivona Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility" in ELSEVIER - Future Generation Computer Systems 25 pg 599-616, 2009

[6] Anju Mohan, Shine, "Survey on Live VM Migration Techniques", International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 1, January 2013, ISSN: 2278 – 1323

[7] Sujesha Sudevalayam and Purushottam Kulkarni, "Affinity-aware modelling of CPU usage with communicating virtual machines" in Journal of Systems and Softwares 86 (2013) 2627-2638.

[8] Pablo Graubner, Matthias Schmidt and Bernd Freisleben, "Energy-Efficient Virtual Machine Consolidation" in IEEE Computer Society, IT Pro, March/April 2013.

[9] Soramichi Akiyama⋆, Takahiro Hirofuchi†, Ryousei Takano† and Shinichi Honiden, "MiyakoDori: A Memory Reusing Mechanism for Dynamic VM Consolidation", 2012 IEEE Fifth International Conference on Cloud Computing

[10] Tiago C. Ferreto, Marco A. S. Netto, Rodrigo N. Calheiros and Cesar A.F. De Rose, "Server Consolidation with migration control for virtualized data centers" in Future Generation Computer Systems 27 (2011) 1027-1034.

[11] Dawei Huang, Deshi Ye, Qinming He, Jianhai Chen, Kejiang Ye, "Virt-LM: A Benchmark for Live Migration of Virtual Machine", March 14–16, 2011, Karlsruhe, Germany. Copyright 2011 ACM 978-1-4503-0519-8/11/03

[12] Mayank Mishra, Anwesha Das, Purushottam Kulkarni and Anirudha Sahoo, "Dynamic Resource Management using Virtual Machine Migrations" in IEEE Communication Magazine (Cloud Computing: Networking and Communication Challenges), September 2012.

[13] Haikum Liu, Hai Jin, Xiaotei Liao, Chen Yu and Cheng-Zhong Xu, "Live Virtual Machine Migration via

Asynchronous Replication and State Synchronisation" in IEEE Transactions on Parallel and Distributed Systems, VOL 22, NO 12, December 2011.
[14] Kejiang Ye, Xiaohong Jiang, Dawei Huang, Jianhai Chen and Bei Wang, "Live Migration of Multiple Virtual Machines with Resource Reservation in Cloud Computing Environments" in IEEE 4th International Conference on Cloud Computing, 978-0-7695-4460-1, 2011.
[15] Febio Checconi, Tomasso Cucunotta and Manuel Stein, "Real Time Issues in Live Migration of Virtual Machines" in Euro-Par'09 Proceedings of the 2009 international conference on Parallel processing (ACM Digital Library), Pages 454-466 [20] Fabio Checconi1, Tommaso Cucinotta1, and Manuel Stein2, "Real-Time Issues in Live Migration of Virtual Machines"
[16] Anton Beloglazov and Rajkumar Buyya, "Managing Overloaded Hosts for Dynamic Consolidation of Virtual Machines in Cloud Data Centers Under Quality of Service Constraints", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS , VOL. 24, NO. 7, JULY 2013
[17] Jyoti Shetty, Anala M R, Shobha G, "A Survey on Techniques of Secure Live Migration of Virtual Machine", International Journal of Computer Applications (0975 – 8887) Volume 39– No.12, February 2012
[18] Zhenhao Pan, Yaozu Dong, Yu Chen Lei Zhang Zhijiao, "CompSC: Live Migration with Pass-through Devices", March 3–4, 2012, London, England, UK.Copyright 2012 ACM 978-1-4503-1175-5/12/03.
[19] Bradley Morgan, "Virtualization",[Published on 4 Jan. 2006 / Last Updated on 24 Jan. 2013
[20] Wei Huang #1, Qi Gao #2, Jiuxing Liu 3 , Dhabaleswar K. Panda, "High Performance Virtual Machine Migration with RDMA over Modern Interconnects".
[21] Daniel A. Menasc´e, "Virtualization: Concepts, Applications, and Performance Modeling", NMA501-03-1-2033.
[22] Dhruv Garg Kamal Kant Abhay Bansal, "Review of Virtual Machine Migration in Datacenters", Volume 3, Issue 6, June 2013 ISSN: 2277 128X International Journal of Advanced Research in Computer Science and Software Engineering
[23] Gandhi Kishan Bipinchandra, Prof. Rajanikanth Aluvalu , Dr.Ajay Shanker Singh, "Intelligent Resource Allocation Technique For Desktop-as-a-Service in Cloud Environment".
[24] Swapnil M. Parikh, "A Survey on Live Virtual Machine Migrations and its Techniques", Computer Engineering and Intelligent Systems, ISSN 2222-1719 (Paper) ISSN 2222-2863, Vol.5, No.3, 2014

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Academic conference: http://www.iiste.org/conference/upcoming-conferences-call-for-paper/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar