

Efficient Resource Allocation for Virtualized Energy Aware Live Migration in Cloud Computing

L. Arulmozhiselvan
Computer Science and Engineering
Easwari Engineering College
Chennai, India

Mrs. N. Senthamarai, M.E,(PH.D)
Computer Science and Engineering
Easwari Engineering College
Chennai, India

Abstract—Cloud computing infrastructure, platform, and software (applications) as services, which are made available to consumers as subscription based services. Cloud computing is defined as the study and practice of designing, manufacturing, using and disposing of computers, servers, and associated sub systems such as monitors, printers, storage devices, and networking and communications systems efficiently and effectively with minimal or no impact on the environment. In this new area of technology, cloud computing offers delivery of on-demand computing resources, everything from application to data center over the Internet on a pay for use basis. Cloud computing offers guaranteed services. Cloud computing allows customers to scale up and down their resources based on their dynamic needs. Cloud computing is quite popular among users for providing on demand services. It refers to both the applications delivered as services over the Internet and the hardware and systems software in the data centers that sustain such services. The proposed work is to make a better use of distributed resources, which the massive computers constitute, while offering dynamic flexible infrastructures and Quality-of-Service (QoS) guaranteed service.

Keywords— Cloud Computing; IaaS; DVFS; Quality-of-Service; Service level Agreement; Virtual machine Scheduling

I. INTRODUCTION

Cloud computing is the major technologies in computer industries. Many companies migrate to these technologies due to reduction in maintenance cost. From the current cloud computing environment there is number of application that consists of millions of module, where these application serve from huge amount of quantity to the users and the user request becomes more dynamic. This Cloud services allow individuals and businesses to the use hardware and software that are managed by the third parties at remote locations. One of the major part in cloud computing is datacenters which provides service all over the world. This is an on demand service because it offers flexible resources allocation, migration and guaranteed services in pay as-you-use manner to public. In the current cloud environment huge amount of energy consumed at the datacenter. This project presented dynamic job migration of virtual machines with the DVFS concept. Also the project satisfies the customer needs by delivering the service that consumes less energy.

According to this the term Migration of job done between virtual machines by considering the energy

consumed and deadline as crucial factors. Moreover, we have integrated our Dynamic Voltage and Frequency Scaling (DVFS) algorithm in CPU utilization model to specify the required frequency for each task depending on the task complexity and the deadline. As well as with the concept of DVFS, total CPU utilization will be found and check for migrate the job to multiple virtual machines or allocate the job to the single virtual machine.

The term DVFS and Green Computing is defined as the environmentally sustainable computing. It refers to the attempts to maximize the use of power consumption and energy efficiency and to minimize the cost and power. The main purpose of green computing is to investigate new computer systems, computing model and applications with the low cost and low power consumption.

1. ENERGY SAVING TECHNIQUES

Where the various level of energy saving techniques is included in this section the following some of the techniques are given below.

A. Concept of cloud computing

Cloud computing is a model where it enables the on demand network to share a configurable computing resources such as application and services.

B. Concept of platform computing

It is based on the service oriented deployment of public and private cloud. Where the Microsoft proposed three solution, window live, window azure and platform cloud.

2. SCHEDULING OF WORK LOAD

The other approach of the term is scheduling the workload, it schedule across the server selected as the function cost to operate in order to improve the efficiency in this approach it is used the queuing theory principle and also proposes the optimal efficiency can be achieved by maintaining to the server.

3. LIVE VIRTUAL MACHINE MIGRATION

In cloud computing, virtualization technology enables VM migration to balance load in the data centers. Basically, migration is done to manage the resources dynamically. It has following goals:

A. Server Consolidation

The main goal of server consolidation is to remove the problem of server sprawl. Server consolidation tries to pack VMs from lightly loaded host on to fewer machines to fulfil resources needs. The free physical machines can be switched off. It will reduce the power consumption and in turn reduce operational cost. This can be achieved by live virtual achieve migration

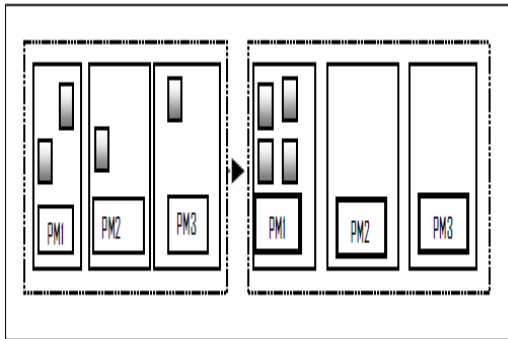


Fig 1: Server Consolidation

B. Load Balancing

Load balancing removes the situation of large difference in resource usage level of the physical machines. It avoids machines from getting overloaded in the presence of low loaded machines. Live migration is used to balance the load across the systems. The whole system load can be balanced by moving virtual machines from fully loaded physical machines to low loaded physical machines.

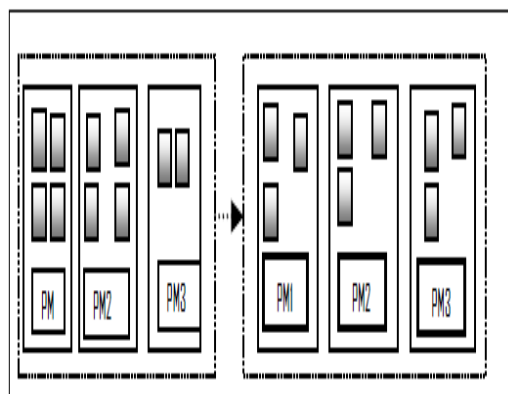


Fig 2: Load Balancing

C. Steps evolved in migration

Live migration means the migration of a VM from source physical machine to destination when the virtual machine is powered up. Virtual machine migration should occurs in such a way that it should minimize both total migration time and down time.

Following logical steps evolved in migration:

Stage 0 Pre-migration: Virtual machine migration is initiated with an active virtual machine on source host A. To speed up future migration process, a destination host is pre-selected and the required resources for migrating VM should be guaranteed.

Stage 1 Reservation: The request is generated to move an operating system from physical host A to physical

host B. First of all it is confirmed that the required resources are present on host B and virtual machine container of same size is reserved. If the required resources are not there then the VM continuously run on host A.

Stage 2 Iterative Pre copy: During first iteration, the whole memory pages are moved from host A to host B. Next iteration only copy those pages which becomes dirty during previous phases.

Stage 3 Stop and Copy: In this phase running operating system instance is suspended at host A and the network traffic is redirected from host A to host B. The remaining memory pages and CPU state are transferred. At the end of stop and copy stage, both host A and host B has consistent suspended copy. The pages at a are considered primary and are resumed if any type of failure occurs.

Stage 4 Commitment: Indication is given by host B to host A that the consistent operating system image is received. Acknowledgement is given by host A and original virtual machine is discarded by host A. Now host B is primary host.

Stage 5 Activation: The migrated virtual machine on host B is activated. Device drivers are reattached to the destination machine and post migration code advertise IP address.

II. SYSTEM ARCHITECTURE

The users will submit the length of the job that is supposed to be submitted to the resource manager. The frequency calculator checks for the possible frequencies that could complete the job within the particular time limit. The objective is to choose the minimum frequency among the available frequencies. Migration checker checks for the load of the entire job, it checks whether the job could be allocated in a single VM totally if it couldn't allocate in a single VM load balancing is done.

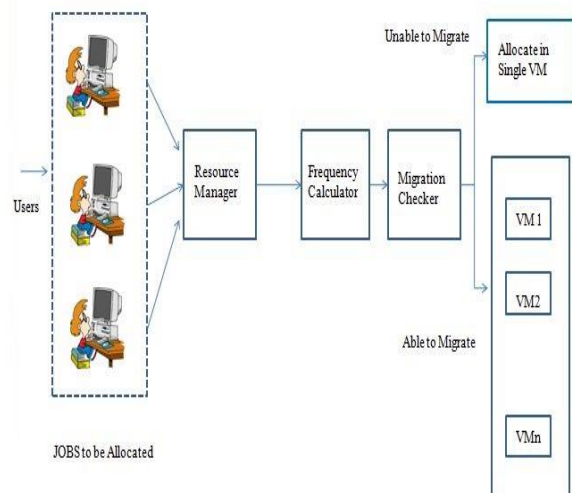


Fig 3: System Architecture

III. PROPOSED SYSTEM

The aim of this proposed system is to assign job to the virtual machine such a manner where it consume the less amount of energy and power saving. where the user will submit the job length and execute the job in a deadline and where the term deadline present the maximum length of the job in execution time for that the algorithm checks the possible frequencies that completes the job in a particular time.

And where now check the job that can migrated from one virtual machine to another virtual machine which consume the less amount of energy. Then the total energy is calculated by taking the power consumption both in static and idle. in such cases, the total energy is consumed after the migration it would be more than that of the without migration .hence this job is migrated from one virtual machine to other virtual machine that will consume less amount of energy. The concept of linear programming is used to find the percentage of job to be migrated to the new virtual machine.

Using the technique in such a manner that the deadline is met and then allocated to two virtual machines. The virtual machine is started one after the other and execution is completed within the deadline.

IV. EXPERIMENTS

In our experiment, we have worked with just one datacenter. We took up with 100 host on this datacenter which in turn is running 100 virtual machines on those hosts. Each node comprises of one CPU core with 10 GB ram/network bandwidth and storage space of 1TB. The host comprises of 2500, 2000, 1000 and 500 MIPS accordingly. For each virtual machine on host ram size is 1024, 512, 256 & 128 MB respectively and bandwidth size is 100 Mbit/s with VM size 2.5 GB. For our experiment we have just worked with one resource. Initially the VMs are considered to be utilized by 100% of time. Firstly, we tried to work on analysis of concept of Adaptive Migration Threshold and its implementation on Cloudsim Toolkit.

The CPU chosen can work in six frequency levels; 1.596GHz, 1.729GHz, 1.862GHz, 1.996GHz, 2.128GHz and 2.261GHz. According to the concept of DVFS, time of execution of the job in each of the frequencies differ from each other. The more the frequency, less is the execution time. In the same manner, the more the frequency, more is the energy consumed. Thus frequency is directly proportional to the energy and inversely proportional to the execution time. The dynamic power consumed by the processor is calculated by the formula:

$$P = Cf^3 \quad (1)$$

Where P is the power consumed, C is the proportional coefficient, f is the dynamic frequency chosen by the governor.

Step 1:

Here the frequency to be taken by the processor is decided based on the actual workload along with the

deadline. There may be more than one frequency that will complete the job in that time limit. The model will choose only the minimum frequency. In our case we assume deadline=3.0. The execution times in each of the frequencies are as given the Even 540 Advancements in Automation and Control Technologies though, the frequencies namely 1.995, 2.128, 2.261 can meet the deadline 3.0, we select the lowest frequency among these three namely 1.995. With this the initial step is over. The total energy consumed by the processor is the cumulative power consumption over the total execution time. Thus energy is calculated by the following formula

$$E = \int_{t=0}^n p(t) \quad (2)$$

Here E is the energy consumed, P is the power consumed at time t and t ranges from 0 to total time of execution. The CPU performance is represented by means of MIPS (Millions of Instructions per Second). In this case E=25.

Step 2:

It is now checked whether the same job can be allocated to more than one virtual machine so that the total energy consumed is less and time of execution falls within the deadline. If the condition is found true, then a percentage of the job is migrated to another virtual machine and both the virtual machines would be executing the job. This is done by the usage of the mathematical concept called Linear Programming. The following formula is used to find the expected time of execution of the job in one virtual machine.

$$t = (\gamma - (\mu * \text{deadline})) / (\alpha - \mu) \quad (3)$$

Here, t is the time, γ is the job length, μ is the MIPS of the virtual machine to which migration is to be taken place and α is the minimum MIPS obtained as a result of Step1. In this case E= 24.07. The screenshots given that the user is giving the input job length = 15000 and deadline = 3.0. The split up of energy to be consumed by both virtual machines, if migrated is also calculated.

V. CONCLUSION

Efficient Resource Allocation has reduced the energy consumption in datacenters by the use of Dynamic Voltage and Frequency Scaling (DVFS) technique along with partial job migration from one virtual machine to another virtual machine under the constraint that migrated job should be completed within the stipulated deadline in an efficient manner. Resource allocation and efficient energy consumption techniques also implemented in cloud datacenter. In the current situation the default governor in DVFS and Live Migration must be described. The proposed work is to make a better use of distributed resources, which the massive computers constitute, while offering dynamic flexible infrastructures and Quality-of-Service (QoS) guaranteed service. In future it will be overcome all the governors and migration can be implement along with other techniques to improve the energy efficiency.

REFERENCES

- [1] Mauro Gaggero and Luca Caviglione, "Predictive Control for Energy Aware Consolidation in Cloud Datacenters" IEEE Transactions on Control Systems Technology, 1063-6536, 2015.
- [2] Christine Mayap kamga, "CPU frequency emulation based on DVFS" IEEE/ACM Fifth International Conference on Utility and Cloud Computing, 367 – 374, 2012.
- [3] Seyed Mohammad Ghoreyshi, "Energy-Efficient Resource Management of Cloud Datacenters Under Fault Tolerance Constraints" International Green Computing Conference, 1-6, 2013.
- [4] Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, Cesar A. F. De Rose, Rajkumar Buyya, "CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource" John Wiley & Sons, Ltd., 23-50, 2010.
- [5] Kyong Hoon Kim, Anton Beloglazov, Rajkumar Buyya, "Power-Aware Provisioning of Cloud Resources For Real-Time Services" ACM, 978-984, 2009.
- [6] Rajkumar Buyya, Anton Beloglazov, "Energy Efficient Resource Management in Virtualized Cloud Data Centers", 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 2010.
- [7] Von Laszewski, G., Wang, L., Younge, A. J., & He, X. "Power-Aware Scheduling of Virtual Machines in DVFS-enabled Clusters". Paper presented at the Proc. of IEEE International Conference on Cluster Computing 2009, New Orleans, LA, USA. (2009).
- [8] D. Meisner, B. T. Gold, and T. F. Wenisch, "PowerNap: Eliminating Server Idle Power", Proc. of ASPLOS, 2009
- [9] Duy, T. V. T., Sato, Y., & Inoguchi, Y. "Performance evaluation of a Green Scheduling Algorithm for energy savings in Cloud computing". Paper presented at the IEEE International Symposium on Parallel & Distributed Processing, 2010.
- [10] R. Bianchini and R. Rajamony, "Power and energy management for server systems," IEEE Computer, vol. 37, no. 11, pp. 68–74, 2004