Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 3, March 2015, pg.226 – 229

SURVEY ARTICLE

A Survey on Energy Aware Scheduling of VMs in Cloud Data Centers

Ms. Ganya Poonacha C[1], Ms. Keerti Priya K[2], Dr. Prakash Sheelvanthmath[3], Prof. Ashwini[4]

Student, MTech (Software Engineering), EPCET, Bangalore, India[1]
Student, MTech (Software Engineering), NHCE, Bangalore, India[2]
Head of Department, Computer Science & Engineering Department, EPCET, Bangalore, India[3]
Assistant Professor, Information Science & Engineering Department, NHCE, Bangalore, Indi4[4]
[1] ganyacp@gmail.com; [2] keertipriya.k@gmail.com

Abstract- ***Cloud Data Centers have become indispensable infrastructure for computing and data storage that facilitates the development diversified services offered by the cloud. It also houses cloud services and cloud based resources which consume large amount of electrical energy and thus emit a lot of CO2 into the surrounding environment. In cloud there is an increased necessity for virtualization and high utilization of the servers and their resources, hence the energy consumed by the data center is increasing drastically. This paper depicts the different techniques involved in scheduling the workload of the servers in order to minimize the energy consumption.***

Keywords: Cloud computing, Virtual machine, Data center

I. INTRODUCTION

In recent years, emerging Cloud Computing concept besides increasing need for computing application has made a great impact on communication and computing systems. Cloud Computing provides an environment for users to access a pool of resources such as Applications, Development Environments, Virtual Machines and so on, by using network communication media to access computing resource. The superiority of this model of access lies in the way it is on demanded fashion and "Pay-per-use" model. Due to its resource sharing nature, this model can cause huge savings in resource consumption cost, maintenance cost, resource upgrading cost, etc. Moreover, it is predicted that Cloud Computing will be dominant in whole IT industry, from user application in form of web services or online storages to whole users machine as Virtual Machines that store on remote computer and user can access to that through a terminal whenever and wherever they wish to. On the other hand, increasing demand on computing causes massive energy consumption and consequently causes pollution to the environment. R. Brown et al. in a report to USA congress in 2006 estimated energy consumption in US IT industry about 4.5 billion USD and predictions show this value will double in 2011, if it is to be continued.

In a research conducted in on more than 5000 production servers over more than six months, they observed that in most of the times servers worked at 11-50% of machines total utilization capacity and it caused huge loss of investigation money, resource utilization, electricity power and consequently increase in total cost. Therefore, this research aims to solve this problem and reduce costs by exploiting this phenomena in Cloud Computing "Infrastructure as a Service" (Iaas) environment where users are provided with their virtual machines on their demands and configuration of the virtual machines determined by an agreement on specifics of virtual machines such as amount of RAM memory, Storage capacity, amount of Bandwidth and CPU power, this agreement made between user and cloud administrator and known as "Service Level Agreement (SLA)". So the main idea is to consolidate virtual machines in cloud data center such that virtual machines with under-utilized resources use resources in share mode and turn off extra machines to reduce energy consumption of host machines and save resources.

II. LITERATURE SURVEY

This section mainly discusses about the papers, websites that are referred while making this dissertation report. All these papers and websites provide information related to learning of collective behavior, their existing solutions, methods used and also their advantages & limitations.

- In, ***"Efficient Resource Management for Cloud Computing Environments,"*** [1] whose authors are A.J. Younge, G. von Laszewski, Lizhe Wang, S.Lopez-Alarcon and W. Carithers describes "Green Cloud Framework" to optimize energy consumption in Cloud Computing Environment. The framework defines computing resource management that are efficient as well as the green computing technologies that can be adapted and also applied to Cloud system. The technique used here is Dynamic Voltage and Frequency Scaling (DVFS) within clusters and Supercomputers [2][3] which reduces the operating frequency and voltage and this will lead to decrease in consumption of power of a given computing resource. The relative aspects of this work consist of VM scheduling according to minimize power consumption by hosts in datacenter and exploit putting multiple VM on a host to reduce number of active hosts. Also considered here is a VM image management that suggest some optimization in order to reduce VM image size and improve VM boot time according to Green Computing aims.
- In, ***"Performance Evaluation of a Green Scheduling Algorithm for Energy Savings in Cloud Computing"*** [4] whose authors are Truong Vinh Truong Duy, Yukinori Sato and Yasushi Inoguchi suggests to put any workload into virtual machines and resource allocate to virtual machines according to requirements of applications and they formulate "Resource allocation problem" to classic "Bin-packing problem" and try to solve that with classic Bin-packing solutions. They divide virtual machines into sub-groups according to their size and used Best-fit algorithm to put them on host machines; they also defined Over-provision ratio value for hosts and tried to find its optimum value.

  Duy and Inoguchi, in an effort, investigated the state transition of host machines in energy optimization process and specified energy consumption of each transition and delay caused by it and tried to find out best transition state. They also suggested the use of Green Scheduling algorithm in integration with a prediction mechanism based on Neural Network to predict workload demands on virtual machines in order to optimize the power consumed by the server in cloud computing environment by shutting down the servers that are unused. This predictor will help the algorithm to make appropriate decisions of turning off/on a server that is being under-utilized [4].

  Benefits: The neural predictor is used for predicting the future load demands on servers based on historical demand.

  Liabilities: However neural networks take much time to train and react to change slowly on resource demands.

- In, ***"Energy Efficient Resource Management in Virtualized Cloud Data Centers,"*** [5] whose authors are Anton Beloglazov and Rajkumar Buyya, introduced a distributed solution to optimize energy efficiency by using the algorithm "Minimization of Migration" as suggested by them. Mainly there are 3 parameters that are considered in virtual machines placement in data-centers and let SLA be violated due to energy saving and redefined problem of energy saving into two-parameter problem with saving energy while keeping value of SLA Violation as low as possible, finally they used "CloudSim" [6] simulator to evaluate their purposed algorithm.

  Benefits - Minimize power consumption and handle strict QoS requirements

  Liabilities - Lacks implementation on a real system and experimental evaluation.

- In, ***"Energy-Aware Resource Allocation Heuristics for Efficient Management of Data Centers for Cloud Computing,"*** [7] whose authors are Anton Beloglazov, Jemal Abawajy, and Rajkumar Buyya, made a survey on existing energy optimization approaches and have divided their optimization area into two sub-phases: "VM Selection" and "VM Placement", and introduced "Modified Best-Fit Decreasing" algorithm for VM Placement sub-phase and three methods "Highest Potential Growth", "Random Choice" and "Minimization of Migration" as VM Selection sub-phase and compared them with simulation results.

- Furthermore, Anton Beloglazov and Rajkumar Buyya in their recent work in ***"Adaptive Threshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Data Centers,"*** [8] introduced new concept of "Adaptive Threshold" for host machines which used statistical prediction based on VM resource usage to determine upper and lower utilization thresholds dynamically according to host`s load and compared new methods with all their old methods. Their works in [5][7][8][6] composed general base of this work.

  Benefits: This approach based on heuristics which allows a reasonable performance even for large-scales. In this the algorithm that is proposed will not ensure the fulfillment of QoS and also it is possible to violate SLA due to the variability in workload. It is possible to dynamically consolidate VM with minimum SLA violation and no of VM migration.

- In, ***"Online Self-Reconfiguration with Performance Guarantee for Energy-Efficient Large-Scale Cloud Computing Data Center,"*** [9] whose authors are Haibo Mi, Gang Yin, Huaimin Wang, Yangfan Zhou, Dianxi Shi, Lin Yuan, tried to solve VM Placement by introducing GABA method, which stands for genetic algorithm based approach. It is an algorithm that adaptively self-reconfigures the virtual machines in large-scale data-centers that are virtualized and consisting of heterogeneous nodes. GABA Algorithm is based on online-reconfiguration algorithm on virtual machines with web server running on them. This algorithm can decide efficiently the optimal number of virtual machines online for each of the application. It will also decide the VM's physical locations according to the dynamic environmental conditions and time varying requirement. They modeled VMs and Hosts as genes and chromosomes and tried to find out best combination of them [9]. In their paper they considered CPU utilization as only virtual machine`s resource.

  Benefits: It allows self-reconfiguration of data-centers without any explicit specifications. It is possible to efficiently perform online searching of optimal solutions through complex and vast possible configuration spaces.

- In ***"pMapper: power and migration cost aware application placement in virtualized systems"*** [10] whose authors are A.Verma, P.Ahuja, and A.Neogi describes the architecture follwed by the implementation of poweraware application placement framework, called pMapper. This includes different scenarios that involve managing of power as well as performance using VMs. pMapper is like a solution provider for the most practical possibility, such as, minimization of power under performance constraints. Here, the used algorithms, framework and the interfaces are being implemented with commercial IBM performance-oriented workload manager which already exists. There are benchmarked Applications on server platforms that are virtualized are benchmarked in order to create a utilization-based power models of the application and server combinations and quantify the cost related to virtualization. This study provides us the knowledge about the structure of the power-aware placement problem which can be used for designing application placement solutions that are tractable. In this paper the migration cost models, power models are used. And also used here is a power-aware placement algorithms for designing dynamic placement controller which, under various assumptions, will performs much better than the static or load balancing placement controller with increasing heterogeneity of the server platforms, i.e. their power models.

  Benefits: It supports virtualization. It brings down the consumption of power as well as loss in performance.

  Liabilities: It does not take into consideration the network bandwidth, the memory, and the advanced application of idle states.

- In, ***"Energy Efficient Allocation of Virtual Machines in Cloud Data Centers"***, [11] whose authors are Anton Beloglazov and Rajkumar Buyya. This paper proposes heuristics for dynamic reallocation of VMs. This is based on the current resources requirements, along with ensuring reliable QoS. The main objective of the reallocation is to reduce the number of physical nodes which are serving the current workload; whereas idle nodes are switched off in order to reduce power consumption. A detailed survey has been done in power efficient resource management in data centers, e.g. [12], [13]. In contrast to previous work, the approach mentioned above can very well handle heterogeneous infrastructure, heterogeneous VMs and strict QoS requirements. The algorithms does not depend on any particular type of workload and also do not require any type of knowledge about applications executing on VMs.

Benefits: The advantages of this paper are Minimizing consumption of power and satisfying performance requirements

Liabilities: Considers CPU only as a resource and not disk storage, network bandwidth as they also consume significant amount of power.

III. CONCLUSION

The data center will continue to expand and evolve over the years, because of which there is always an urge for optimizing the energy. This survey paper shows the different techniques involved in scheduling the workload of the servers in order to minimize the energy consumption. This paper also mentions advantages and disadvantages of each paper.

REFERENCES

[1] A.J. Younge, G. von Laszewski, Lizhe Wang, S. Lopez-Alarcon, and W. Carithers, "*Efficient Resource Management for Cloud Computing Environments*," in International Green Computing Conference, 2010, pp. 357 - 364.

[2] W. chun Feng and C.-H. Hsu and "A Feasibility Analysis of Power Awareness in Commodity-Based High-Performance Clusters," in CLUSTER, 2005, pp. 1–10.

[3] W. Feng and C. Hsu and, "A power-aware run-time system for highperformance computing," in Proceedings of the 2005 ACM/IEEE conference on Supercomputing. IEEE Computer Society Washington, DC, USA, 2005.

[4] Yukinori Sato, Truong Vinh Truong Duy and Yasushi Inoguchi, "Performance Evaluation of a Green Scheduling Algorithm for Energy Savings in Cloud Computing," in *IEEE International Symposium on Digital Object Identifier*, 2010, pp. 1-8.

[5] Rajkumar Buyya and Anton Beloglazov, "Energy Efficient Resource Management in Virtualized Cloud Data Centers," in *10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, 2010.

[6] Rodrigo N Calheiros, Anton Beloglazov, Rajiv Ranjan, Rajkumar Buyya, and César A. F De Rose, *CloudSim:A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms*. USA: Wiley Press, 2010.

[7] Anton Beloglazov, Rajkumar Buyya and Jemal Abawajy, "Energy-Aware Resource Allocation Heuristics for Efficient Management of Data Centers for Cloud Computing," *journal of Future Generation Computer Systems*, vol. 27, no. 8, April 2011.

[8] Rajkumar Buyya and Anton Beloglazov and "Adaptive Threshold-Based Approach for Energy-Effcient Consolidation of Virtual Machines in Cloud Data Centers," in *ACM/IFIP/USENIX 11th International Middleware Conference*, 2010.

[9] Huaimin Wang, Haibo Mi, Yangfan Zhou, Dianxi Shi, Gang Yin and Lin Yuan, "Online Self-Reconfiguration with Performance Guarantee for Energy-Efficient Large-Scale Cloud Computing Data Centers," in *IEEE International Conference on Services Computing*, 2010, pp. 514 - 521.

[10] A. Verma, A. Neogi, and P. Ahuja ―pMapper: power and migration cost aware application placement in virtualized systems, in Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware. Springer-Verlag New York, Inc., 2008, pp. 243–264.

[11] Anton Beloglazov and Rajkumar Buyya - Energy Efficient Allocation of Virtual Machines in Cloud Data Centers in 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 2010.

[12] D. Kusic, N. Kandasamy, J. O. Kephart, J. E. Hanson, and G. Jiang, "Power and performance management of virtualized computing environments via lookahead control," Cluster Computing, vol. 12, no. 1, pp. 1–15, 2009.

[13] S. Srikantaiah, F. Zhao and A. Kansal, "Energy aware consolidation for cloud computing," Cluster Computing, vol. 12, pp. 1–15, 2009.

