Green Cloud Computing: Efficient Energy-Aware and Dynamic Resources Management in Data Centers

Sara DIOUANI, Hicham MEDROMI

Engineering research laboratory (LRI)
System Architecture Team
ENSEM, Hassan II University of Casablanca, Morocco

Abstract—**The users of Cloud computing over the last years are constantly increasing since it has become a very important technology in the computing landscape. It provides client decentralized services and a pay-as-you-go model for consuming resources. The growing need for the cloud services oblige the providers to adopt an enlarged sized data center infrastructure which runs thousands of hosts and servers to store and process data. As a result, these large servers engender a lot of heat with visual carbon emission in the air, as well as important energy consumption and higher operating cost. This is why researches in energy economics continue to progress including energy saving techniques in servers, in the network, cooling, and renewable energies, etc. In this paper, we tackled the existing energy efficient methods in the green cloud computing fields and we put forward our green cloud solution for data center dynamic resource management. Our proposed approach aims to reduce the infrastructure energy consumption and maintain the required performances.**

Keywords—*Cloud computing; green cloud; data center; energy consumption; resource management*

I. INTRODUCTION

The recent revolution in information and communication technologies, despite all its benefits for our way of life, has reinforced our dependence on energy. More than 3 billion people use electronic terminals on a daily basis, the electricity consumption of which is strongly linked to the time of use, which is constantly increasing [1]. To support this revolution, huge data centers have sprung up all over the world. Service providers and data hosts are competing for investment in these data centers, real information factories for which profits cannot be counted. However, the power consumption of data centers is a new puzzle that the scientific communities are trying to solve with enormous difficulties.

The issue of energy has emerged in recent years as a central concern that humanity faces with great urgency. Population growth, the gradual exhaustion of previously explored resources, and more recently the rise of information technologies have turned this issue into a challenge that researchers and industry players are tackling [2].

In the field of information technology and in 2010, about 1.5% of the world's electricity has been consumed by computer centers [3]. This is steadily increasing due to the evolution of many domains and especially cloud computing.

In recent years, new problems have emerged in view of the environmental considerations that are increasingly present in our society. In 2014, the power consumption of Data centers exceeded 42 TWh and by 2020 the resulting CO2 production will reach 670 million metric tons annually [4].

Also, so as to face the peak demand of coming requests, servers in the data center, are constantly over-provisioned in working state, which generates wasting a large amount of energy [5]. One of the options to reduce the power consumption of data centers is to reduce the number of idle servers or to switch them into low-power sleep states as part of the green cloud IT vision [6].

Our work is to explore new ways to improve energy efficiency in cloud data centers. Specifically, the goal is to dynamically optimize energy consumption in cloud computing data centers by optimizing the use of its resources by adopting various policies for host and virtual machine (VM) overload detection, migration VM selection and VM placement policies.

The remainder of this paper is arranged as follows. In Section II, a literature review on the energy-aware solution on green cloud computing is presented. In Section III, our solution is proposed. Finally, the conclusions and future works are drawn in Section IV.

II. RELATED WORK ON GREEN CLOUD COMPUTING

Generally, the green cloud data centers are related to three principal methods which are: dynamic voltage frequency scaling (DVFS), the scheduling using renewable energy, and dynamic power management (DPM) [7].

So as to have an estimation of the energy consumed by the cloud client application, researchers in [8] have performed a power model under various DVFS policies. In [9], the flow pattern of the cloud tasks is studied and according to the results obtained, the researchers attempt to adjust the incoming VM tasks with in-demand frequency using DVFS.

In [10], the DVFS is adopted to minimize the power consumed in mobile cloud task scheduling, except that this approach does not consider the On and Off control of the servers. In the study [11] the three approaches including DVFS, request dispatching, and dynamic service management are joined so as to reduce the energy consumption. Yet, the limitation is that these researchers have admitted that the servers providing different services are active all the time.

In the SaaS cloud platform, an analytical framework which monitors the states of VMs (idle/busy) is developed. It characterizes and optimizes the power performance tradeoff

[12]. In another search [13] only one type of sleep mode with shutdown are used in a method for reducing the energy consumption, except that this solution is insufficient in the case of quick responses. In [14], they turned off a physical host in cloud computing by the adoption of two parameters which are time and load, so as to save energy consumption. If a VM had its working time exceeds a threshold fixed time, this VM will be displaced. And if a physical host had a load less than a fixed threshold, it will be turned off.

In this research [15], they relied on the processor workload, the disk workload and the ratio of performance degradation as a metrics of a VM so as to reallocate it. In [16], they used an algorithm based on CPU and memory of the VM and the server as parameters to consolidate them. They select VMs which are underloaded regarding the server. Some researchers have come to prove that this server consolidation approach is not adapted for large data centers. Also, multiple meta-heuristic algorithms of optimization are used for consolidating VMs such as those cited in [17]-[19].

According to many previous studies [20], [21], the green cloud computing gathers the efficient management of cloud resources and reducing energy consumption while assuring the quality of service requirements in the service level agreements SLA.

A system view of the green cloud computing where reducing energy is a must can be shown as in Fig. 1. It illustrates that when the cloud user sends an application request, the resource manager controls the resource utilization, and is in charge of the allocation of VMs in physical hosts while ensuring the respect of the SLA. Also, the underloaded servers are turned off while some others may be switched on for consolidating physical hosts as needed (see Fig. 1).

This survey helps to realize that existing limitations in the previous researches and which all of the major energy parameters (e.g. CPU, memory and so forth) necessary to ensure an ideal energy efficiency are not taken into consideration.
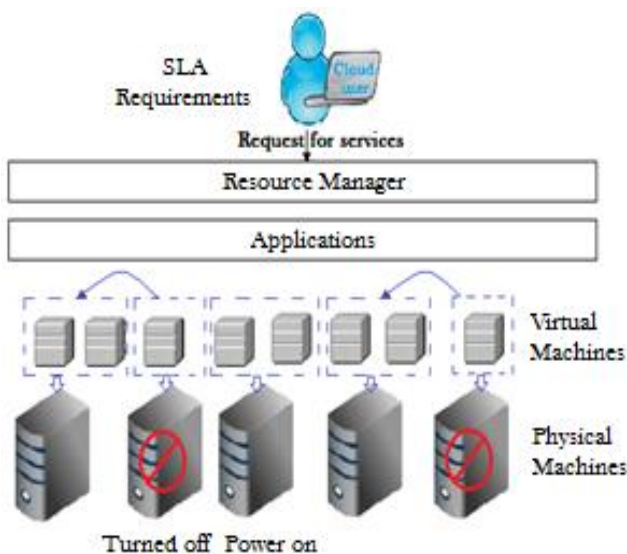

Fig. 1. Overview of green cloud computing.

III. DYNAMIC AND ENERGY-AWARE SOLUTION FOR RESOURCES MANAGEMENT

Virtual machine consolidation techniques are a means to improve energy efficiency and the utilization of cloud data center resources. However, aggressive VM consolidation approaches lead to physical host over-utilization and generate massive undesired VM migrations, which cause degradation of the performance of both the hosts and the VM [22].

Additionally, it has been a significant challenge to improve energy efficiency and resource utilization in the data center while delivering services with guaranteed quality of service (QoS).

To address this problem, we propose an enhancing energy-efficient and QoS dynamic resource management method, which consists of four principal modules. Our solution considers all major parameters related to the efficiency of the cloud data center energy.

In fact, the relevant energy parameters include the CPUs, the memory amount, the disk storage space, the quantity of transmitted message in the network (bandwidth) and the available amount of input/output operations per second (IOPS) on the physical host.

Besides, the VM placement depends on some precise Service Level Agreement restrictions as follows:

- The affinity constraint: between VMs aimed to get an optimal placement by considering the requirement that two VMs for example, must be placed on the same physical machine (PM). This condition refers to interdependent VMs that use jointly data with each other in prespecified deadlines.

- The security constraint: may be, for example, detaching two VMs on different servers (or even two data centers), so as to ensure their separation.

- The migration constraint: means that it is possible to execute the VM placement only on a set of precisely stated machines, or to maintain a VM on the same host (or same data center).

We defined also other energy parameters which are : the total number of VMs placed on a PM, the total number of PMs used, the number of reallocations of a VM, the period of time of VM interruption in the migration period, the percentage of maximum and minimum use of VMs/PMs and the response time of a task at the level of a VM (SLA).

The aspect of the sustainability of data is also taken into consideration and which refers to ensuring in real time the replication of each data to multiple hosts (such as a primary and backup host).

Our solution uses an infrastructure model that is informed of the state of the system at any moment desired. It is focused on dynamically manage VM allocation and displacement in the data center, in terms of performance, availability system, cost and instantly energy consumption. This approach adopts optimized resource allocation and live migration through a decision and analysis mechanism and with an effective respect of the strict SLA requirements.
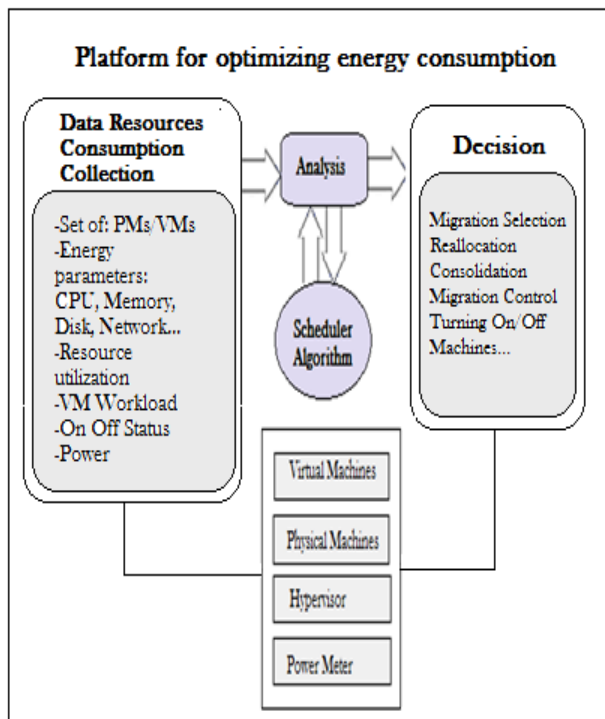
Fig. 2. Overview of the proposed solution.

As Fig. 2 explains, the overview of the proposed optimized cloud model platform and which includes different management stages as following: collecting monitoring data, exploiting these data to calculate a better resources placement, draw a plan to reallocate and dynamically manage resources, and applying the proposed actions (see Fig. 2).

The collection component maintains a periodic monitoring and collection of data related to workload, electrical consumption and the use of resources from the cluster's PMs. These data include CPU, disk storage, memory usage, etc. Also, the electricity collection is done by using power consumption measurement tools (for instance Power Meter or Wattmeter).

The analysis module adopts a built-in scheduler algorithm to perform decisions and which evaluates the results of resource data use. The designed and implemented algorithm does not rely on a particular type of workload and does not need any information about the applications running on the VMs. Its execution involves exact power parameters and high-resolution measurements.

The input of the algorithm is the collection of each PM with its allocated VMs, the different collected resource information, and the specified energy consumption parameters so as to determine which resource to allocate and where to displace it. A new optimized resources placement plan with a set of nodes underloaded and overloaded (to deactivate, turn on later ...) is the output of the algorithm execution.

The decision component aims for the migration, reallocation, and consolidation of the resources. Also, resources can be turned on/off according to the instructions of

the analysis model and based on the results previously obtained from the execution of the scheduler algorithm.

IV. CONCLUSION AND FUTURE WORKS

Today's IT services are using the cloud computing solutions so as to offer to its clients the required services efficiently. Except that the high use of the cloud engenders a large growth in its data center infrastructure. In this case, unfortunately, an enormous amount of electrical energy is consumed and a high amount of carbon dioxide is emitted in the air.

Thus, reducing the energy consumption in cloud data centers while assuring an optimized management of its resources including VMs and servers is becoming a needful aim to achieve. This requirement is related to the green cloud concept by which we can contribute to the environmental protection.

In this paper, various techniques for enhancing the green cloud resources allocation are discussed and which are based essentially on virtualization, migration, and consolidation. Thus, the proposed solution provides an optimized resource management while considering all major energy parameters and major possible constraints of VMs allocation in PMs and which influences on the energy consumed in the cloud computing data center. Also, we focused on taking energy-performance trade-off in concern.

In future, we will detail and implement the scheduler algorithm while respecting the defined Service Level Agreements and the required Quality of services.

REFERENCES

[1] « Internet Society Global Internet Report 2015 », p. 142.

[2] A. E. I. Ahmed and S.-N. Nada, "Integrated Framework For Green ICT: Energy Efficiency by Using Effective Metric and Efficient Techniques For Green Data Centres".

[3] P. Corcoran, A. Andrae, "Emerging trends in electricity consumption for consumer ICT", 2013.

[4] S. F. Smith, "Is Scheduling a Solved Problem?", in Multidisciplinary Scheduling: Theory and Applications, Springer, Boston, MA, 2005, p. 3‑17.

[5] G. S. Akula and A. Potluri, "Heuristics for migration with consolidation of ensembles of Virtual Machines" in 2014 Sixth International Conference on Communication Systems and Networks (COMSNETS), 2014, p. 1‑4.

[6] S. Agarwal, A. Datta and A. Nath, "Impact of green computing in IT industry to make eco friendly environment," Journal of Global Research in Computer Science, vol. 5, no. 4, pp. 5–10, Apr. 2014.

[7] C. Gu, Z. Li, H. Huang, and X. Jia, "Energy Efficient Scheduling of Servers with Multi-Sleep Modes for Cloud Data Center", IEEE Trans. Cloud Comput., p. 1‑1, 2018.

[8] F. D. Rossi, M. Storch, I. de Oliveira, and C. A. F. D. Rose, "Modeling power consumption for DVFS policies" in 2015 IEEE International Symposium on Circuits and Systems (ISCAS), 2015, p. 1879‑1882.

[9] A. P. Florence, V. Shanthi, and C. B. S. Simon, "Energy Conservation Using Dynamic Voltage Frequency Scaling for Computational Cloud", ScientificWorldJournal, vol. 2016, p. 9328070, 2016.

[10] X. Lin, Y. Wang, Q. Xie, and M. Pedram, "Task Scheduling with Dynamic Voltage and Frequency Scaling for Energy Minimization in the Mobile Cloud Computing Environment", IEEE Trans. Serv. Comput., vol. 8, no 2, p. 175‑186, mars 2015.

[11] Y. Chen, C. Lin, J. Huang, X. Xiang, and X. (Shen, "Energy Efficient Scheduling and Management for Large-Scale Services Computing

Systems", IEEE Trans. Serv. Comput., vol. 10, no 2, p. 217‑230, mars 2017.

[12] Z. Zhou, F. Liu, H. Jin, B. Li, B. Li, and H. Jiang, "On arbitrating the power-performance tradeoff in SaaS clouds" in 2013 Proceedings IEEE INFOCOM, 2013, p. 872‑880.

[13] V. K. M. Raj and R. Shriram, "A study on server Sleep state transition to reduce power consumption in a virtualized server cluster environment" in 2012 Fourth International Conference on Communication Systems and Networks (COMSNETS 2012), 2012, p. 1‑6.

[14] C. C. Lin, P. Liu, and J. J. Wu, "Energy-efficient Virtual Machine Provision Algorithms for Cloud Systems" in 2011 Fourth IEEE International Conference on Utility and Cloud Computing, 2011, p. 81‑88.

[15] M. Sharifi, H. Salimi, and M. Najafzadeh, "Power-efficient distributed scheduling of virtual machines using workload-aware consolidation techniques", J. Supercomput., vol. 61, no 1, p. 46‑66, juill. 2012.

[16] A. Murtazaev and S. Oh, "Sercon: Server Consolidation Algorithm using Live Migration of Virtual Machines for Green Computing", IETE Tech. Rev., vol. 28, no 3, p. 212‑231, mai 2011.

[17] N. Quang-Hung, P. D. Nien, N. H. Nam, N. H. Tuong, and N. Thoai, "A Genetic Algorithm for Power-Aware Virtual Machine Allocation in Private Cloud" in Information and Communication Technology, Springer, Berlin, Heidelberg, 2013, p. 183‑191.

[18] X.-D. Zuo and H.-M. Jia, "An energy saving heuristic algorithm based on consolidation of virtual machines" in 2013 International Conference on Machine Learning and Cybernetics, 2013, vol. 04, p. 1578‑1583.

[19] D. A. Alboaneen, H. Tianfield, and Y. Zhang, "Glowworm Swarm Optimisation Algorithm for Virtual Machine Placement in Cloud Computing" in 2016 Intl IEEE Conferences on Ubiquitous Intelligence Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), 2016, p. 808‑814.

[20] J. Huang, K. Wu, and M. Moh, "Dynamic Virtual Machine migration algorithms using enhanced energy consumption model for green cloud data centers" in 2014 International Conference on High Performance Computing Simulation (HPCS), 2014, p. 902‑910.

[21] M. Shaden and A. Heba, "Review of Energy Reduction Techniques for Green Cloud Computing", International Journal of Advanced Computer Science and Applications, 2016:184-195.

[22] J. Sahoo, S. Mohapatra, and R. Lath, "Virtualization: A Survey on Concepts, Taxonomy and Associated Security Issues" in 2010 Second International Conference on Computer and Network Technology, 2010, p. 222‑226.