A Survey on Energy Aware Scheduling in Virtualized Clouds

Pavitha N[1*] & S. N. Shelke[2]

1 Department of Computer Engineering, Sinhgad Academy of Engineering, Pune,Maharashtra, India
Email: pavithanrai@gmail.com

[2]Assistant Professor, Sinhgad Academy of Engineering, Pune, Maharashtra, India

Abstract:

Cloud computing has recently received considerable attention, as a promising approach for delivering ICT services by improving the utilization of data centre resources. Energy safeguarding is a most important concern in cloud computing systems because it can bring several important benefits such as reducing operating costs, increasing system reliability, and prompting environmental protection. Meanwhile, power-aware scheduling approach is a promising way to achieve that goal. At the same time, many real-time applications, e.g., signal processing, scientific computing have been deployed in clouds. These applications necessitate an Energy-Aware Scheduling in Virtualized Clouds. In principle, cloud computing can be an inherently energy-efficient technology for ICT provided that its potential for significant energy savings that have so far focused on hardware aspects, can be fully explored with respect to system operation and networking aspects. In this backdrop this paper formalizes various energy aware scheduling strategies for virtualized clouds.

Keywords:

Cloud Computing, data centre, virtualized cloud, energy aware scheduling

I. Introduction

With the significant advances in Information and Communications Technology (ICT) over the last half century, there is an increasingly perceived vision that computing will one day be the 5th utility (after water, electricity, gas, and telephony). This computing utility, like all other four existing utilities, will provide the basic level of computing service that is considered essential to meet the everyday needs of the general community. To deliver this vision, a number of computing paradigms have been proposed, of which the latest one is known as Cloud computing. [1]

Cloud Computing promises reliable services delivered through next-generation data centers that are built on virtualized compute and storage technologies. Consumers will be able to access applications and data from a "Cloud" anywhere in the world on demand. The consumers are assured that the Cloud infrastructure is very robust and will always be available at any time. Computing services need to be highly reliable, scalable, and autonomic to support ubiquitous access, dynamic discovery and composability. In particular, consumers indicate the required service level through Quality of Service (QoS) parameters, which are noted in SLAs established with providers. Of all these paradigms, the recently emerged Cloud computing paradigm appears to be the most promising one to leverage and build on the developments from other paradigms. [1]

"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."[1]

Cloud computing offers utility-oriented IT services to users worldwide. Based on a pay-as-you-go model, it enables hosting of pervasive applications from consumer, scientific, and business domains. However, data centers hosting Cloud applications consume huge amounts of electrical energy, contributing to high operational costs and carbon footprints to the environment. Therefore, we need Green Cloud computing solutions that can not only minimize operational costs but also reduce the environmental impact. [3]

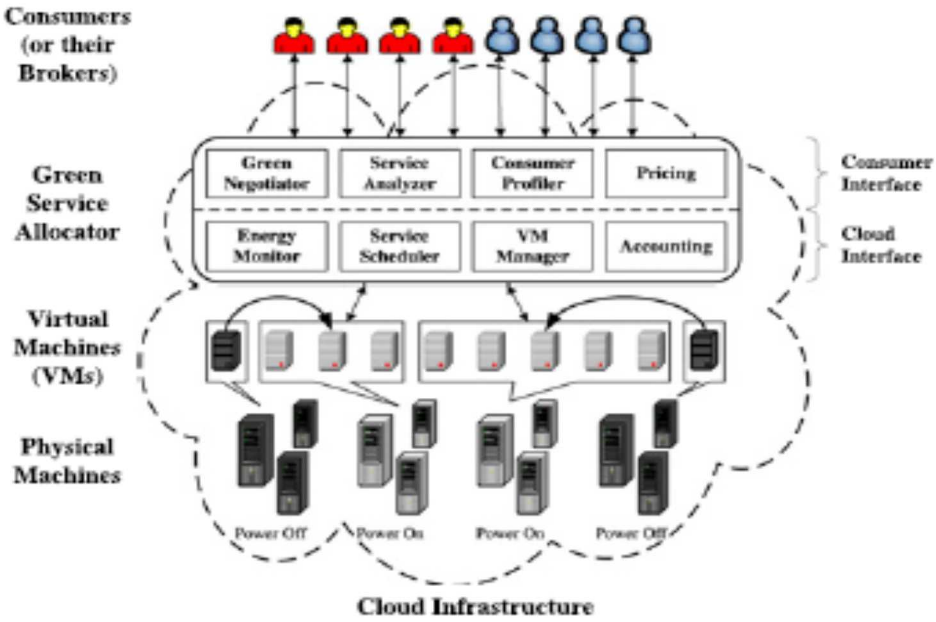

Figure: Green Cloud Architecture

In a broad sense, scheduling algorithms can be classified into two categories: static scheduling and dynamic scheduling [10]. Static scheduling algorithms make scheduling decisions before tasks are submitted, and are often applied to schedule periodic tasks [11]. However, aperiodic tasks whose arrival times are not known a priori must be handled by dynamic scheduling algorithms. [2]

II. Energy-efficient scheduling algorithms for virtualized clouds

Nowadays, virtualization technology has become an essential tool to provide resource flexibly for each user and to isolate security and stability issues from other users [4]. Therefore, an increasing number of data centers employ the virtualization technology when managing resources. Correspondingly, many energy-efficient scheduling algorithms for virtualized clouds were designed. [2]

A. GreenCloud: A New Architecture for Green Data Centre

Liu et al. aimed to reduce energy consumption in virtualized data centers by supporting virtual machine migration and VM placement optimization while reducing the human intervention [5].

B. A Dynamic Configuration Model for Power-Efficient Virtualized Server Clusters

Petrucci et al. presented the use of virtualization for consolidation and proposed a dynamic configuration method that takes into account the cost of turning on or off servers to optimize energy management in virtualized server clusters [6].

C. Dynamic Provisioning Modeling for Virtualized Multi-Tier Applications in Cloud Data Center

Bi et al. suggested a dynamic resource provisioning technique for cluster-based virtualized multitier applications. In their approach, a hybrid queuing model was employed to determine the number of VMs at each tier [7].

D. Power and Migration Cost Aware Application Placement in Virtualized Systems

Verma et al. formulated the power-aware dynamic placement of applications in virtualized heterogeneous systems as continuous optimization, i.e., at each time frame, the VMs placement is optimized to minimize energy consumption and to maximize performance [8].

E. Energy-Aware Resource Allocation Heuristics for Efficient Management of Data Centers for Cloud Computing

Beloglazov et al. proposed some heuristics for dynamic adaption of VM allocation at runtime based on the current utilization of resources by applying live migration, switching idle nodes to sleep mode [3].

F. Energy-Efficient and Multifaceted Resource Management for Profit-Driven Virtualized Data Centers

Goiri et al. presented an energy-efficient and multifaceted scheduling policy, modeling and managing a virtualized data centre, in which the allocation of VMs is based on multiple facets to optimize the provider's profit [9].

G. An Adaptive Model-Free Resource and Power Management Approach for Multi-Tier Cloud Environments

Wang et al. investigated adaptive model-free approaches for resource allocation and energy management under time-varying workloads and heterogeneous multitier applications, and multiple metrics including throughput, rejection amount, queuing state were considered to design resource adjustment schemes [12].

H. Energy-Efficient Management of Virtual Machines in Eucalyptus

Graubner et al. proposed an energy-efficient scheduling algorithm that was based on performing live migrations of virtual machines to save energy, and the energy costs of live migrations including pre- and post-processing phases were considered [13].

I. Real-Time Tasks Oriented Energy-Aware Scheduling in Virtualized Clouds

Xiaomin Zhu et al. proposed the energy-efficient scheduling by rolling-horizon optimization to efficiently guarantee the schedulability of real time tasks and at the same time striving to save energy by dynamic VMs consolidation.[2]

III. Conclusion

Energy safeguarding is a most important concern in cloud computing systems because it can bring several important benefits such as reducing operating costs, increasing system reliability, and prompting environmental protection. Meanwhile, power-aware scheduling approach is a promising way to achieve that goal. At the same time, many real-time applications, e.g., signal processing, scientific computing have been deployed in clouds. These applications necessitate an Energy-Aware Scheduling in Virtualized Clouds. In this locale this paper formalized various energy aware scheduling strategies for virtualized clouds.

References

[1] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, “Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the Fifth Utility,” Future Generation Computer Systems, vol. 25, no. 6, pp. 599-616, 2009.

[2] Xiaomin Zhu, Laurence T. Yang, Huangke Chen, Ji Wang, Shu Yin, Xiaocheng Liu "Real-Time Tasks Oriented Energy-Aware Scheduling in Virtualized Clouds", IEEE TRANSACTIONS ON CLOUD COMPUTING, VOL. 2, NO. 2, APRIL-JUNE 2014, pp. 168- 180.

[3] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-Aware Resource Allocation Heuristics for Efficient Management of Data Centers for Cloud Computing," Future Generation Computer Systems, vol. 28, pp. 755-768, 2012.

[4] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the Clouds: A Berkeley View of Cloud Computing,"Technical Report UCB/EECS-2009-28 Univ. of California, Berkeley, 2009.

[5] L. Liu, H. Wang, X. Liu, X. Jin, W. He, Q. Wang, and Y. Chen, "GreenCloud: A New Architecture for Green Data Center," Proc. Sixth Int'l Conf. High Performance Distributed Computing (HPDC '08), pp. 29-38, June 2008.

[6] V. Petrucci, O. Loques, and D. Moss_e, "A Dynamic Configuration Model for Power-Efficient Virtualized Server Clusters," Proc. 11th Brazillian Workshop Real-Time and Embedded Systems (WTR '09), May 2009.

[7] J. Bi, Z. Zhu, R. Tian, and Q. Wang, "Dynamic Provisioning Modeling for Virtualized Multi-Tier Applications in Cloud Data Center," Proc. Third IEEE Int'l Conf. Autonomic Computing (ICAC '06), pp. 15-24, June 2006.

[8] A. Verma, P. Ahuja, and A. Neogi, "pMapper: Power and Migration Cost Aware Application Placement in Virtualized Systems," Proc. Ninth ACM/IFIP/USENIX Int'l Conf. Middleware (Middleware '08), pp. 243-264, Dec. 2008.

[9] I. Goiri, J.L. Berral, J.O. Fit_o, F. Juli_a, R. Nou, J. Guitart, R. Gavald_a, and J. Torres, "Energy-Efficient and Multifaceted Resource Management for Profit-Driven Virtualized Data Centers," Future

Generation Computer Systems, vol. 28, pp. 718-731, 2012.

[10] Y.-K. Kwok and I. Ahmad, “Static Scheduling Algorithms for Allocating Directed Task Graphs to Multiprocessors,” ACM Computation Survey, vol. 31, no. 4, pp. 406-471, 1999.

[11] X. Qin and H. Jiang, “A Novel Fault-Tolerant Scheduling Algorithm for Precedence Constrained Tasks in Real-Time Heterogeneous Systems,” J. Parallel Computing, vol. 32, no. 5, pp. 331-356, 2006.

[12] X. Wang, Z. Du, and Yi Chen, “An Adaptive Model-Free Resource and Power Management Approach for Multi-Tier Cloud Environments,” The J. Systems and Software, vol. 85, pp. 1135-1146, 2012.

[13] P. Graubner, M. Schmidt, and B. Freisleben, “Energy-Efficient Management of Virtual Machines in Eucalyptus,” Proc. IEEE Fourth Int’l Conf. Cloud Computing (CLOUD ’11), pp. 243-250, 2011.