ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016

ENERGY CONSERVATION THROUGH EFFICIENT GREEN CONTROL ALGORITHM IN CLOUD

Ms.K.Preethi, Mrs.M.Jeyasheela rakkini
ME CSE
Assistant professor,
Department of computer science and engineering,
Mookambigai College of engineering,
Pudukkottai,
e-mail:kpreethi255@gmail.com
e-mail:jayasheelarakkini@yahoo.in

Abstract— *Cloud computing is a new paradigm for delivering remote computing resources through a network. To achieving energy efficiency control and simultaneously satisfying a performance guarantee have become critical issues for cloud providers. Energy conservation is a major concern in cloud computing systems because it can bring several important benefits such as reducing operating costs, increasing system reliability, and prompting environmental protection. An efficient green control algorithm is first proposed for solving constrained optimization problems and making costs, performances tradeoffs in systems with different power saving policies. The proposed power saving policies combined with Efficient Green Control algorithm can effectively reduce cost, increase response time and provide high arrival rates.*

Keywords*: Cost optimization, Energy-efficiency control, Response time, Power saving policy*

I. INTRODUCTION

Cloud computing is a new service model for sharing a pool of computing resources that can be rapidly accessed based on a converged infrastructure. In the past, an individual use or company can only use their own servers to manage application programs or store data. Nowadays, resources provided by cloud allow users to get on demand access with minimal management effort based on their needs. Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) are all existing service models. For example, an Amazon web service is a well-known IaaS that lets users perform computations on the Elastic Compute Cloud (EC2).

To satisfy uncertain workloads and to be highly available for users anywhere at any time, resource over-provisioning is a common situation in a cloud system. However, most electricity-dependent facilities will inevitably suffer from idle times or low utilization for some days or months since there usually have off-seasons caused by the nature of random arrivals. In fact, servers are only busy 10-30 percent of the time on average .As cloud computing is predicted to grow, substantial power consumption will result in not only huge operational cost but also tremendous amount of carbon dioxide (CO_2) emissions. Therefore, an energy efficient control, especially in mitigating server idle power has become a critical concern in designing a modern green cloud system. Ideally, shutting down servers when they are left idle during low-load periods is one of the most direct ways to reduce power consumption. Unfortunately, some negative effects are caused under improper system controls. First, burst arrivals may experience latency or be unable to access services. Second, there has a power consumption overhead caused by awakening servers from a power-off state too frequently. Third, the worst case is violating a service level agreement (SLA) due to the fact that shutting down servers may sacrifice quality of service (QoS).The SLA is known as an agreement in which QoS is a critical part of negotiation. A penalty is given when a cloud provider violates performance guarantees in a SLA contract. In short, reducing power consumption in a cloud system has raised several concerns, without violating the SLA constraint or causing additional power consumption are both important. To avoid switching too often, a control approach called N policy, had been extensively adopted in a variety of fields, such as computer systems, communication networks, wireless multimedia, etc. Queuing systems with the N policy will turn a server on only when items in a queue is greater than or equal to a predetermined N threshold, instead of activating a power-off server immediately upon an item arrival.

Three power-saving policies that (a) switching a server alternately between idle and sleep modes, (b) allowing a server repeat sleep periods and (c) letting a server stay in a sleep mode only once in an operation cycle are all considered for comparison. The main objective is to mitigate or eliminate unnecessary idle power consumption without sacrificing performances. The challenges of controlling the service rate and applying the N policy to minimize power consumption and simultaneously meet a response time guarantee are first studied. To address the conflict issue between performances and power-saving, a tradeoff

All Rights Reserved © 2016 IJARTET

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016

between power consumption cost and system congestion cost is conducted. An efficient green control (EGC) algorithm is proposed to optimize the decision-making in service rates and mode-switching within a response time guarantee by solving constrained optimization problems. As compared to a typical system without applying the EGC algorithm, more cost-saving and response time improvements can be achieved.

This paper contributes to investigate an essential tradeoff between power consumption costs and system performances by applying different power saving policies. To the best our knowledge, applying N- policy for optimizing the mode switching control and simultaneously achieving the minimum cost under a performance guarantee.

II. POWER MANAGEMENT IN CLOUD

A distributed service system consists of lots of physical servers, virtual machines and a job dispatcher. The job dispatcher in our designed system is used to identify an arrival job request and forward it to a corresponding VM manager that can meet its specific requirements. When there has no job in a waiting queue or no job is being processed, a server becomes idle and it remains until a subsequent job has arrived. Generally, a server operates alternately between a busy mode and an idle mode for a system with random job arrivals in a cloud environment.

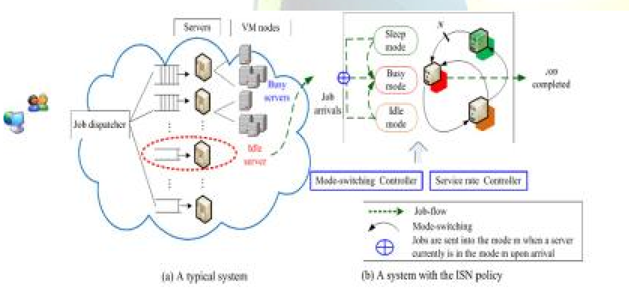


Fig.1 System Architecture

A busy mode indicates that jobs are processed by a server running in one or more of its VMs and an idle mode indicates that a server remains active but no job is being processed at that time. To mitigate or eliminate idle power wasted, three power-saving policies with different energy efficient controls, decision processes and operating modes are presented. First, we try to make an energy-efficient control in a system with three operating modes m ={Busy, Idle, Sleep}, where a sleep mode would be responsible for saving power consumption. A server is allowed to stay in an idle mode for a short time when there has no job in the
System, rather than switch abruptly into a sleep mode right away when the system becomes empty .An idle mode is the only operating mode that connects to a sleep mode. A server doesn't end its sleep mode even if a job has arrived it begins to work only when the number of jobs in a queue is more than the controlled N value.

1. ISN Policy

The job dispatcher in the designed system is used to identify an arrival job request and forward it to a queue of a corresponding VM manager that can satisfy its QoS levels, meet its target web application or specific requirements. When there has no job in a queue or no job is being processed, a server becomes idle and it remains until a subsequent job has been sent to its processor node. Generally, a server operates alternately between a busy mode and an idle mode for a system with random job arrivals in a cloud environment.

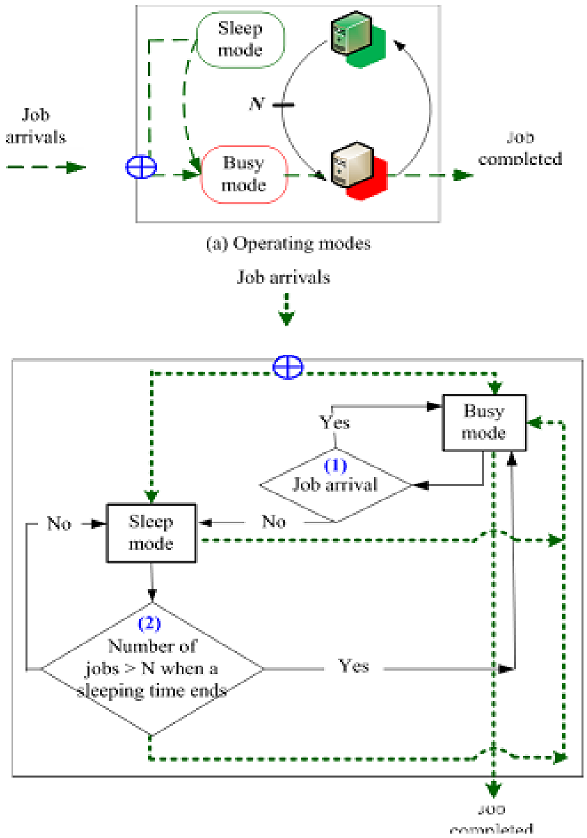
A server is allowed to stay in an idle mode for a short time when there has no job in a queue or no job is being processed, rather than switching abruptly into a sleep mode right away when a system becomes empty. An idle mode is the only operating mode that connects to a sleep mode. A server doesn't end its sleep mode even if there has a job arrival it begins to work only when the number of jobs in a queue is more than the controlled N value.

2. SN and SI Policies

To greatly reduce power consumption, non idle mode operating is considered in the SN policy. It only holds busy and sleep operating modes. Instead of entering into an idle mode, a server immediately switches into a sleep mode when system becomes empty. Similarly, a server switches into a busy mode depending on the number of jobs in the queue to avoid switching too often, the switching restriction is denoted by N. A server switches into a sleep mode immediately when no job is in the system. A server stays in a sleep mode if the number of jobs in the queue is less than the N value. Otherwise, a server switches into a busy mode and begins to work.

All Rights Reserved © 2016 IJARTET

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016



(a) Operating modes

(b) Decision processes

SI policy is similar to SN policy but a server only says in a sleep mode for a given time. When a sleeping time, it will enter into an idle mode or a busy mode depending upon whether a job has arrived or not. According to the switching process from sleep to idle have called such approach SI policy. A server immediately switches into a sleep mode instead of an idle mode when there has no job in the system. A server can stay in a sleep mode for a given time in an operation period. If there has no job in the queue when a sleeping time expires, a server will enter into an idle mode. Otherwise, it switches into a busy mode without any restriction and begins to work. The SN policy having exponential service times, Erlang-k service times and general service times. The SI policy having exponential distributions and deterministic idle times. The response time guarantee is one of the most important performance concerns in designing a green cloud system and no customer wants to suffer from long delay caused by power conservation. Therefore, the SLA constraint is focused on the response time guarantee by considering both the queuing delay and job execution time. The sleep probabilities and the idle probabilities will be reduced as the arrival rate increases since a server has more probability to work stay in a busy mode. Service rates are controlled at higher values with power saving policies and the idle times can be reduced by switching into sleep modes.

III EFFICIENT GREEN CONTROL ALGORITHM

An efficient green control algorithm is proposed to optimize the decision making in service rates and mode switching with in a response time guarantee by solving constraint optimization problems. As compared to typical system without applying the EGC algorithm, more cost saving and response times improvement can be achieved. The main objective of the algorithm is to mitigate or eliminate unnecessary idle power consumption without sacrificing performances. The different powers saving policies are:

i. Illustrate the relationship between the mode switching restriction and traffic load intensity on power consumption cost and system congestion cost.
ii. Examine the idle and sleep probability distributions under different service rates.
iii. Compare the response times and total operational costs with a typical system, where it does not have any energy efficient control.

For an idle server in a data center, power waste is compound by not only the server itself, but also the power distribution losses and air conditioning power usage, which increase power consumption requirements. Most of the idle times reduced or eliminated by switching into sleep modes for systems with power saving policies. An efficient green control algorithm is presented to solve the non linear constrained optimization problem effectively. The sleep probabilities and idle probability will be reduced as the arrival rate increases since a server has more probability to work and stay in busy mode. It is also noted that sleep probabilities are obvious larger than the idle probability in a system with the general policy due to the fact that the proposed algorithm tries to enhance efficiency in busy modes. Therefore, service rates are controlled at higher values with power-saving policies and their idle times can be reduced by switching into sleep modes. Conversely, the general policy focuses only on a performance guarantee and reduces the service rate as low as possible for the purpose of saving operational cost.

The input of the efficient green control algorithm contains an arrival rate, upper bound of the server rate and the waiting buffer, cost parameters, and system parameters used by the ISN policy, SN policy and SI policy. On the output of algorithm find the current service rate, current parameter, and calculate the response time, system utilization. If the current test parameters satisfy the constraint, record the current joint values, and identify it as the approved joint parameters. All parameters have been done, bring the cost parameters into the objective function. Finally joint the parameter values and obtain the minimum cost value. The proposed power saving policies can effectively reduce cost, especially when the arrival rate is low.

All Rights Reserved © 2016 IJARTET

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 20, April 2016

Finally, it measures the performance and cost improvement ratios, which calculate the relative value of improvements to the original value instead of an absolute value. The proposed algorithm allows cloud providers to optimize the decision making in service rate and mode switching restriction, so as to minimize the operational cost without sacrificing a SLA constraint. As compared to a general policy, the benefits of cost savings and response time improvement can be verified.

IV CONCLUSION

The growing crisis in power shortages has brought a concern in existing and future cloud system designs. To mitigate or eliminate unnecessary idle power consumption, three power-saving policies with different decision processes and mode-switching controls are considered. The issue of choosing the most suitable policy among diverse power management policies to reach a relatively high effectiveness has been examined based on the variations of arrival rates and incurred costs. Experimental results show that a system with the SI policy can reach a greater cost-effectiveness than other policies when there has a lower startup cost. It also can significantly improve the response time in a low arrival rate situation. On the other hand, applying others policies can obtain more benefits in the converse situation.

Our proposed algorithm allows cloud providers to optimize the decision-making in service rate and mode-switching restriction, so as to minimize the operational cost without sacrificing a SLA constraint. As compared to a general policy, the benefits of cost savings and response time improvement can be verified. In future, we plan to analyze more key factors that will influence the power consumption and system performances in cloud environments. Looking into finite user populations, traffic load controls, etc. will be another direction of extension to achieve energy-efficiency through a comprehensive system control.

REFERENCES

[1]Ahmed Amokrane, Mohamed Faten Zhani, Rami langar, and Raouf boutaba, “Greenhead:Virtual Data Center Embedding across Distributed infrastrutures,” IEEE Trans. Cloud Computing., vol. 1, no. 1, Jan. 2013.

[2] Guazzone, C. Anglano and M. Canonico, “Energy-efficient resource management for cloud computing infrastructures,” in Proc. IEEE Int. Conf. Cloud Comput. Technol. Sci., 2011, pp. 424–431.

[3]Hetal Joshiyara, “Energy Efficient Provisioning of Virtual Machines in Cloud System using DVFS ,” Indian Journal of Research vol:3,Issue:4,May 2013,ISSN-2250-1991.

[4]Larumbe, and B. Sanso, “A tabu search algorithm for the location of data centers and software components in green cloud computing networks,” IEEE Trans. Cloud Comput., vol. 1, no. 1, pp. 22–35, Jan.–Jun. 2013.

[5]Manasa H.B, Anirban Basu, “ Energy Aware Resource Allocation in Cloud Datacenter”, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-5, June 2013.

[6]M. Mazzucco, D. Dyachuky, and R. Detersy, “Maximizing Cloud Providers Revenues via Energy Aware Allocation Policies,” in Proc. IEEE 3rd Int. Conf. Cloud Comput., 2010, pp. 131–138.

[7]Meisner, B. T. Gold, and T. F. Wenisch, “PowerNap: Eliminating server idle power,” in Proc. 14th Int. Conf. Archit. Support Program. Lang. Oper. Syst., 2009, pp. 205–216.

[8]Q. Zhang, M. Zhani, R. Boutaba, and J. Hellerstein, “Dynamic heterogeneity-aware resource provisioning in the cloud,” IEEE Trans.Cloud Comput., vol. 2, no. 1, pp. 14–28, Jan.–Mar. 2014.

[9]R. Ranjan, L. Zhao, X. Wu, A. Liu, A. Quiroz, and M. Parashar,“Peer-to-peer cloud provisioning: Service discovery and load-balancing,”in Cloud Computing. London, U.K.: Springer, 2010,pp. 195–217.

[10] R. N. Calheiros, R. Ranjan, and R. Buyya, “Virtual machine provisioning based on analytical performance and QoS in cloud computing environments,” in Proc. Int. Conf. Parallel Process., 2011,pp. 295–304.

All Rights Reserved © 2016 IJARTET