ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

An Efficient Green Control Algorithm in Cloud Computing for Cost Optimization

K.Karthika, C.Dhanabal

M. E Student, Dept. of Computer Science and Engineering, Mount Zion College of Engineering and Technology

Pudukkottai, India

Dept. of Computer Science and Engineering, Mount Zion College of Engineering and Technology

Pudukkottai, India

ABSTRACT: Cloud computing is a new paradigm for delivering remotecomputing resources through a network. However, achieving an energy efficiency control and simultaneously satisfying a performance guarantee have become critical issues for cloud providers. In this paper, three power-saving policies are implemented in cloud systems to mitigate server idle power. The challenges of controlling service rates and applying the N-policy to optimize operational cost within a performance guarantee are first studied. A cost function has been developed in which the costs of power consumption, system congestion and startup are all taken into consideration.The effect of energy-efficiency controls on response times, operating modes and incurred cost are all demonstrated. Our objectives are to find the optimal service rates and mode-switching restriction, so as to minimize cost within a response time guarantee under varying arrival rates. An efficient green control algorithm is first proposed for solving constrained optimization problems and making cost/performances tradeoffs in systems with different power saving policies. Simulation results show that the benefit of reducing operational costs and improving response times can be verified by applying the power saving policies are combined with the proposed algorithm as compared to a typical system under a same performance guarantee.

KEYWORDS: cost optimization, energy- efficiency control, response time, power saving policies.

1. 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 this past, an individual use or company can only use their own servers to manage application programs or store data. Now a days resources provided by cloud allow users to get on demand access with minimal management effort based on their needs. Infrastructure as s service (IaaS), platform as a service (PaaS) and software as a service (SaaS) are all existing service models. For example, Amazon web services is a well-known IaaS that lets user's platform computations on the Elastic Compute Cloud (Ec2).

To satisfy uncertain workloads and to be highly available for the users anywhere at any time, over provisioning is a common situation in a cloud system. The main idea of the system toMitigate or eliminate idle power wasted, three power-savingpolicies with different energy efficient controls, decision processes and operating modes are presented.
Software as a service is a software licensing and delivery model in which software is a licensed on a subscription basis and is centrally hosted. It is sometimes referred to as "on-demand software". Three power-saving policies that [1] switching a server alternately between idle and sleep modes, [2] allowing a server repeat sleep periods and [3]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 optimize power consumption and simultaneously meet a response time guarantee are first studied. To address the conflict issue between performances and power-saving, a trade-off between power consumption cost and system congestion cost is conducted.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

We provide a scalable and efficient distributed service integrity attestation framework for large-scale cloud computing infrastructures. The present a novel integrated service integrity attestation scheme that can achieve higher pinpointing accuracy than previous techniques. We describe a result auto-correction technique that can automatically correct the corrupted results produced by malicious attackers. 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.
Burst arrivals may experience latency or be unable to access services. There has a power consumptionoverhead caused by awakening servers from a power-off state too frequently. The worst case is violating a servicelevel agreement (SLA) due to the fact that shuttingdown servers may sacrifice quality of service (QoS). The SLA is known as an agreement in which QoS is a criticalpart of negotiation. A penalty is given when a cloud providerviolates performance guarantees in a SLA contract. Reducing power consumption in a cloud system has raised several concerns, without violating the SLA constraint or causing additional power consumption are bothimportant.

II. RELATED WORK

2.1 Power-Saving in Virtual Machine

The virtual machine placement problem with a goal of minimizing the total energy consumption. A multi-dimensional space partition model and a virtual machine placement algorithm were presented.When a new VM placement task arrived, their algorithm checked the posterior resource usage state for each feasible PM, and then chose the most suitable PM according to their proposed model to reduce the number of running PMs. Considered the problem of providing power budgeting support while dealing with many problems that arose when budgets virtualized systems. They managed power from a VM-centric point of view, where the goal was to be aware of global utility trade-offs between different virtual machines (and their applications) when maintaining power constraints for the physical hardware on which they ran. Their approach to VM-aware power budgeting used multiple distributed managers integrated into the virtual power management (VPM) framework.
Two issues in energy conservation algorithm were addressed-the placement of virtual machine image and the characteristics of virtual machines. Despite that the dynamic programming could find the optimal solution, its time complexity was prohibitive in practice. Theenergy efficiency from the performance perspective wasstudied. Presented a virtual machine based energyefficientdata centers architecture for cloud computing. Then,they investigated the potential performance overheadscaused by server consolidation and lived migration of virtualmachine technology. The potential performances overheadsof server consolidation were evaluated.

2.2 Power-Saving in Computing Infrastructure

The Datacentre Energy Management project was focused on modelling energy consumption in data centers, with a goal to optimize electricity consumption. Their project was focused on collecting data to define basic fuel consumption curves. The objectives were to optimize the network performance, the CO2 emissions, the capital expenditures, and the operational expenditures. The objective of cloud computing was to minimize the power consumption of the network. Their proposed model allowed planners to evaluate different solutions and to make variations in the optimization priorities. Although power management in cloud has attracted considerable research attention, few studies focused on effectively reducing idle server power draw. Unlike previous studies, our paper contributes to investigate an essential trade-off between power consumption costs and system performances by applying different power-saving policies. To the best of our knowledge, applying the N-policy for optimizing the mode-switching control and simultaneously achieving the minimum cost under a performance guarantee has not been considered before.

A. Problem of Virtual machine

a. Not fully achieved for power consumption in cloud based storages.
b. Multi User request in same time not properly response.
c. More emulsion of Carbon di oxide gas due to use of server maintenance.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

III. PROPOSED SYSTEM

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. 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.

B. Solution of propoesd system

a. Proper response to the entire requested user in the cloud computing process.
b. Avoid the over emulsion of Carbon di oxide gas due to use of server maintenance.
c. Practically implemented algorithm and approaches.

3.1Architecture of Green Control Algorithm

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. An EGC algorithm is presented to solve the nonlinear constrained optimization Problem effectively. Meeting a SLA constraint has the highest priority, followed by cost minimization in deciding the optimal solution.

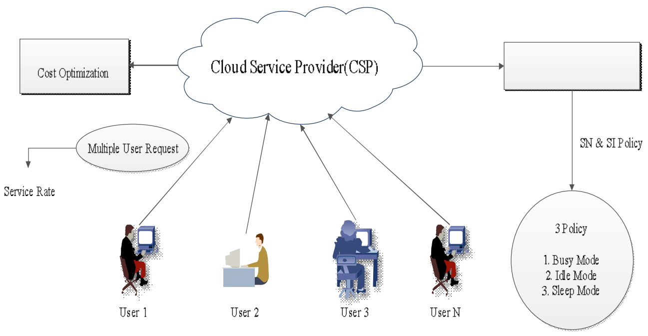

Figure 3.1.1 EGC Architecture

Input for this algorithm Arrival Rate, Cost Parameters, Response Guarantee, System parameters used by three different policies. Processes are Calculate Service, Rates Calculate System Utilization, and Estimate cost value. Output is efficient response time for SLA certification, Effective measurement of cost used by cloud server.

IV. IMPLEMENTATION

4.1 Generation of Queuing Model

Generally Queue maintains several job request given by the authorized users.The job request arrivals follow a Poisson process with parameter and they are served in order of their arrivals, that is, the queue discipline is the first come first served (FCFS).There may have some job requests that need to be performed serially at multiple service stages. Then, applying phase-type distributions allow us to consider a more general situation.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

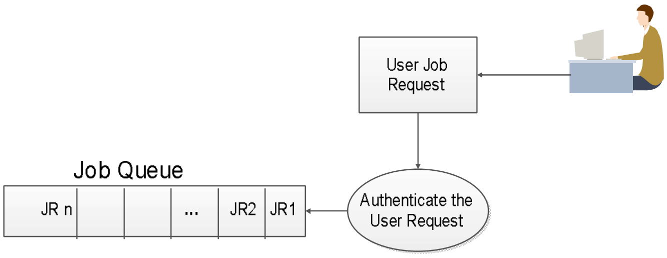

Figure 4.1.1 generation of queuing model

4.2 Implementation of ISN policy

Mainly used when the response needs immediately along some extra energy used. 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. 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.

Two cases starting Busy Mode:

a. Starting a busy mode when a job arrives in an idle mode.
b. Starting a busy mode if the number of jobs in a waiting queue is more than the N value when a sleep period expires.

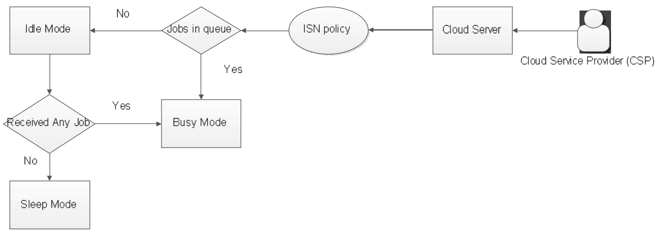

Figure 4.2.1 Implementation of ISN policy

4.3 modelling the SN policy

For a cloud provider who focuses on reducing cost, implementing the SN policy is a better choice to deal with a wide range of arrival rates. According to the switching process (directly to Sleep) and the energy-efficient control (N policy), we have called such an approach the "SN policy. 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. The other policy is designed with no mode-switching restriction and performed under the other energy-efficient control. A server switches into a sleep mode right away rather than an idle mode when there has no job in the system. This is similar to the SN policy but a server only stays in a sleep mode for a given time. When a sleeping time expires, it will enter into an idle mode or a busy mode depending upon whether a job has arrived or not.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

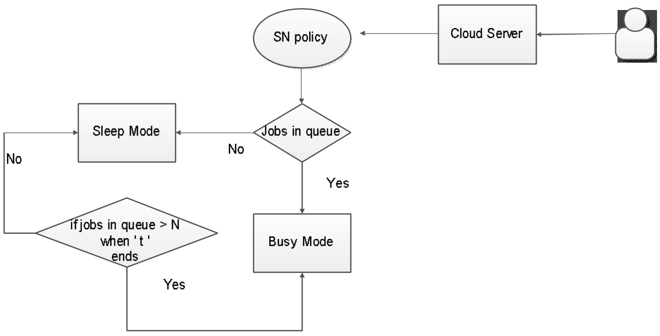

Figure 4.3.1 modelling the SN policy

4.4 Enhance SI policy

It can significantly improve the response time in a low arrival rate situation. According to the switching process (from Sleep to Idle), we have called such an approach "SI policy". The step-by-step decision processes and job flows of the SI policy. A server switches into a sleep mode immediately 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 arrival 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. Let Pmn denote the steady-state probabilities at state(n, m), then the following notations are usedP0n -Probability that there have n jobs in the system when a server is in a busy mode. P1 -Probability that there has no job in the system when aserver is in an idle mode.

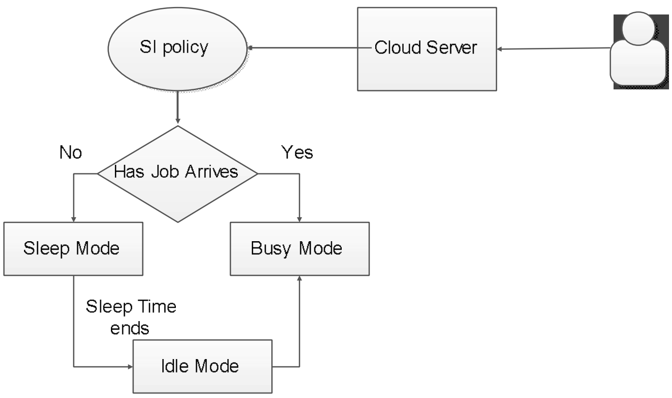

Figure 4.4.1 Enhance SI policy

4.5 Formulation of EGC algorithm

An EGC algorithm is presented to solve the nonlinear constrained optimization problem effectively. Meeting a SLA constraint has the highest priority, followed by cost minimization in deciding the optimal solution (M*; N*). This 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.The objective of the system is power saving and also supporting green computing. Benefits of reducing operational costs and improving response times can be verified by applying the power-saving policies combined with the proposed algorithm as compared to a

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

typical system under a same performance guarantee. Three power-saving policies are implemented in cloud systems to mitigate server idle power.

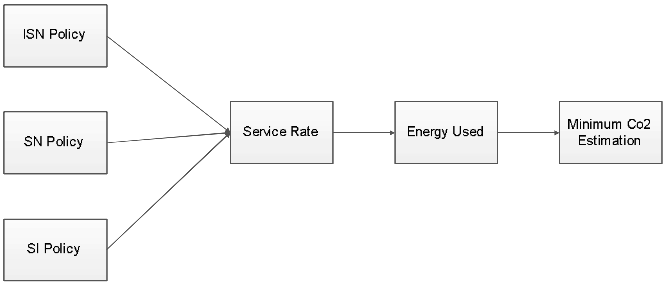

Figure 4.5.1Formulation of ECG algorithm

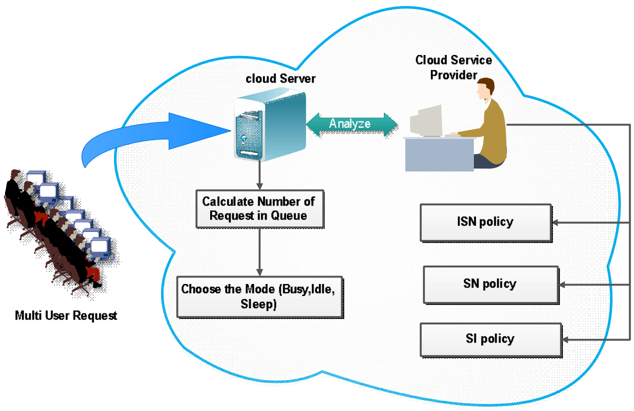

Figure 4.5.2 EGC policies

V. PERFORMANCE ANALYSIS

Three operating modes m = {Busy, Idle, Sleep}, where a sleep
Mode would be responsible for saving power consumption. An idle mode is the only operating mode that connects to a sleep mode. This is the ISN policy performance analysis graphical representation.

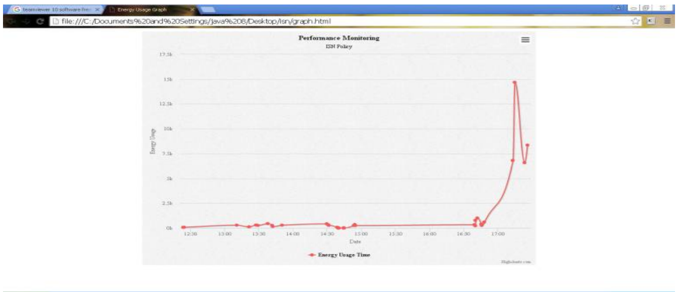

Figure 5.1 ISN Policy Analysis

Directly to sleep if there is having job request it will be going to be normal mode. This is the SN policy performance analysis graphical representation. 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.

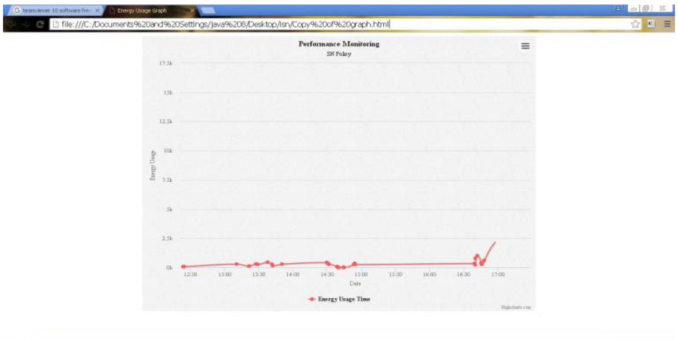

Figure 5.2 SN Policy Analysis

A cloud service provider when use SI policy there has no job it will going to sleep mode. When sleep time ends modes changed sleep to idle mode.

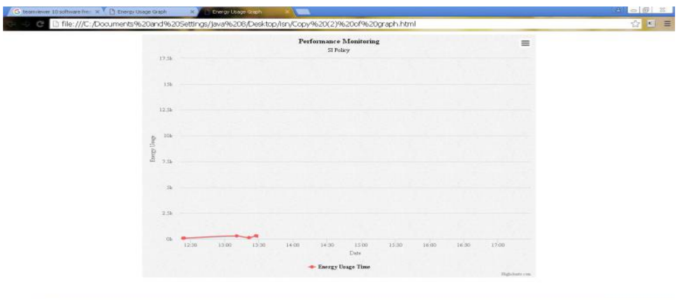

Figure 5.3 SI Policy Analysis

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

The arrival rate of the three policies in the probability

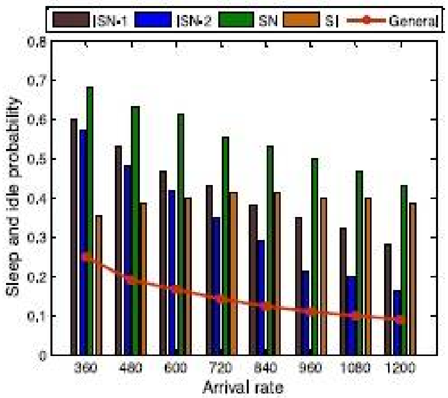

Figure 5.4 Policies arrival rate Analysis

VI. CONCLUSION

To mitigate unnecessary idle power consumption, three power saving policies with different decision processes and mode switching controls are considered.The issue of choosing a suitable policy amongdiverse power managements to reach a relatively high effectivenesshas been examined based on the variations ofarrival rates and incurred costs. Experimental results showthat a system with the SI policy can significantly improvethe response time in a low arrival rate situation. On the

Other hand, applying others policies can obtain more costbenefits when the start-up cost is high. As compared to ageneral policy, cost savings and response time improvementcan be verified. A proposed algorithm ECG 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. 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. That can be resolved by ISN policy.

REFERENCES

[1] Y. C. Lee and A. Y. Zomaya, **"Energy efficient utilization of resources in cloudComputingsystems,"**J.Supercomput, vol. 60, no. 2, pp. 268–280, 2009.

[2] Amokrane, M. Zhani, R. Langar, R. Boutaba, and G. Pujolle, **"Greenhead: Virtual datacenter embedding across distributed infrastructures,"** IEEE Trans. CloudComputing. Vol. 1, no. 1, pp. 36– 49, Jan.–Jun. 2010.

[3] Q. Zhang, M. Zhani, R. Boutaba, and J. Hellerstein,**"Dynamic heterogeneity- awareResource Provisioning In the cloud,"** IEEE Trans. Cloud Comput., vol. 2, no. 1, pp. 14–28, Jan.–Mar. 2012.

[4] D. Meisner, B. T. Gold, and T. F. Wenisch, **"PowerNap: Eliminating server idle power,"** in Proc. 14th Int. Conf. Archit. Support Program. Lang. Oper. Syst., pp. 205–216, 2013.

[5] H. AbdelSalam, K. Maly, R. Mukkamala, M. Zubair, and D. Kaminsky, **"Towards energyefficient change Management in a cloud computing environment,"** in Proc. 3rd Int.Conf. Auton. Infrastructure, Manage. Security: Scalability Netw. Serv., 2009, pp. 161–166, 2015.

[6] J. S. Yang, P. Liu, and J. J. Wu**, "WorkloadCharacteristics-Aware virtual machine consolidation Algorithms,**" in Proc.IEEE 4th Int. Conf. Cloud Comput.Technol. Sci., 2012, pp. 2–49.

[7] A. Beloglazov, R. Buyya, Y. C. Lee, and A. Zomaya, **"A taxonomy and survey of energy-efficient data centers and cloud computing systems,**" Adv. Comput., vol. 82, pp. 47–111, 2011.