IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

An Efficient & Scalable Framework to Reduce Cost of Operation using Vitality Utilization Improvement in Cloud Computing

Banashankari[1], Mr. Chandan Raj. B. R[2]

M. Tech Student, Dept. of CSE, EWIT, Bengaluru, Karnataka, India[1]

Assistant Professor, Dept. of CSE, EWIT, Bengaluru, Karnataka, India[2]

Abstract: Cloud computing paradigm has become one of the most popular platforms for sharing of various remote computing resources in a distributed manner. The existing research trends shows that achieving cost effectiveness considering energy optimization, efficient performance analysis and low computation associated with internal operations turned out to be very challenging for cloud providers which gained a lot of attention from researchers. This project aims to develop a two power saving strategies which can be utilized further to achieve cost effectiveness in terms of energy optimization and low computation time. The proposed study adopts the conventional N-policy strategy for optimizing the operational cost associated with the cloud client-server architecture. The performance evaluation of the proposed prototype ensures further reusability and reutilization to simplify the future research direction.

Keywords: Cloud computing, Cost of operation, Energy optimization, N-policy strategy.

I. INTRODUCTION

Cloud computing is a large-scale distributed computing paradigm in which a pool of computing resources is available to cloud consumers via the Internet. Computing resources, such as processing power, storage, software, and network bandwidth, are represented to cloud consumers as the available public utility services. Infrastructure as- a-Service (IaaS) is a computational service model widely used in the cloud computing paradigm. Here in this method, virtualization technologies can be used to supply resources to cloud customers. The clients can specify the required software stack, e.g., operating systems and applications; then enclose them all together into virtual machines (VMs). The hardware requirement of VMs can also be adjusted by the consumers. At last, these VMs will be outsourced to host in computing environments. With the reservation plan, the cloud consumers will previously reserve the resources in advance.

There may occur under provisioning problem when the reserved resources are unable to fully meet the demand due to its uncertainty. Although this problem can be solved by ordering extra resources by the help of on-demand plan to fit the extra demand, the high cost will be incurred due to more expensive price of resource provisioning with on demand plan. At the same time, the over provisioning problem can occur when the reserved resources are more than the real demand in which part of a resource pool will be underutilized. The main part of the cloud consumer is to minimize the total cost of resource provisioning by reducing the on-demand cost and oversubscribed cost of under provisioning and over provisioning. Therefore, an energy efficient control, particularly in moderating server unmoving force has turned into a basic worry in outlining a cutting edge green cloud framework. In a perfect world, closing down servers when they are left sit without moving amid low-stack periods is a standout amongst the most direct approaches to diminish power utilization. Shockingly, some negative impacts are brought about under despicable framework controls. This study plans to build up a proficient and adaptable structure for planning costs/exhibitions tradeoffs in frameworks with various force sparing approaches.

As to queuing frameworks with two periods of administration in [1]concentrated on the circulated frameworks where all clients holding up in the line get clump administration in the primary period of administration took after by individual administration in the second stage. In [2] presented the two stage queueing framework with server get-aways. In [3] explored the two stage framework with N-policy. The server new companies compare to the preliminary work of the server before beginning the administration. In some real circumstances, the server regularly requires a startup time before beginning every administration period. The first proposed the N-arrangement M/M/1 queueing framework with exponential startup time. A few creators have examined queueing models with server breakdowns and excursions in various structures in later past. Hence the present project aims to propose the system to formulate an efficient and scalable framework to optimize the operational cost in cloud computing servers.

It provides an efficient and scalable framework to reduce cost of operation using energy consumption optimization in cloud computing and also calculated some important parameters like idle probability, sleep probability, busy probability, mean queue length, response time, power consumption cost, system congestion cost and total cost for

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

the sever system. Section I briefly describes about an efficient and scalable framework to reduce cost of operation using energy consumption optimization in cloud computing. Section II gives the background of this topic. Section II deals with literature survey which explains about the previous work done. Section IV explains the design methodology for the proposed method as well as implementation details. Section V gives results and discussion are obtained in this work. Finally, Overall conclusions from this paper and possible extensions are presented in Section VI.

II. BACKGROUND

In most queuing frameworks the server might be subjected to protracted and eccentric breakdowns while serving a client. Case in point, in assembling frameworks the machine might breakdown because of glitch or occupation related issues; in PC frameworks, the machine might be liable to planned reinforcements and anticipated disappointments. In these frameworks, server breakdown results in a time of server inaccessibility time until it is settled. Along these lines, it is important to perceive how the breakdowns influence the level of execution of the framework. Fig.1 shows the block diagram of the proposed system.
Distributed computing is another worldview for conveying remote processing assets through a system. In any case, accomplishing a vitality productivity control and all the while fulfilling an execution ensure have ended up basic issues for cloud suppliers. In this project, the main contributions are summarized as follows.

1. 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 rest mode just once in an operation cycle are all considered for examination. The principle goal is to alleviate or dispose of pointless unmoving force utilization without giving up exhibitions.
2. The difficulties of controlling the administration rate and applying the N-strategy to minimize power utilization and at the same time meet a reaction time assurance are initially concentrated on. To address the contention issue in the middle of exhibitions and force sparing, a tradeoff between force utilization expense and framework blockage expense is directed.
3. A productive green control (EGC) calculation is proposed to advance the choice making in administration rates and mode-solving so as to exchange inside of a reaction time ensure obliged improvement issues. When contrasted with an average framework without applying the EGC calculation, more cost-sharing and reaction time enhancements can be accomplished.

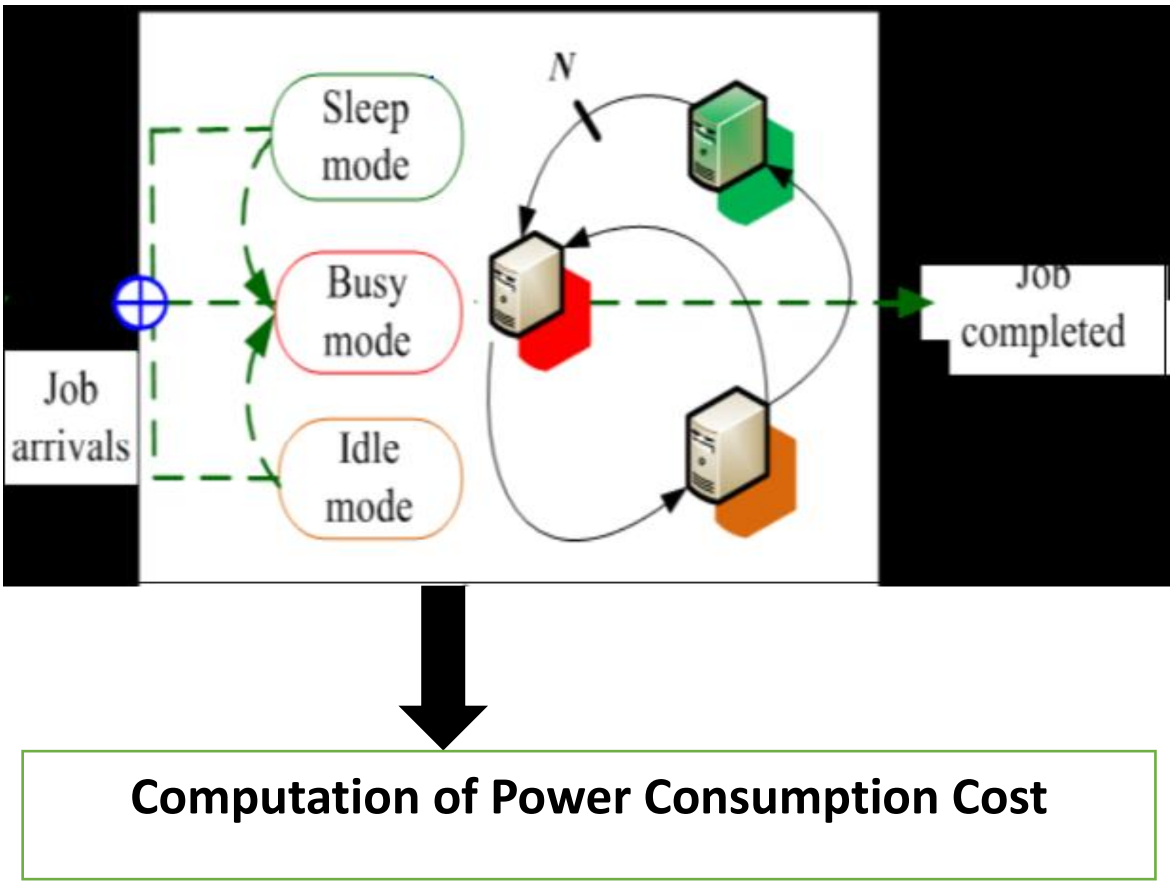

Fig.1 Block diagram of the Proposed Model

DOI 10.17148/IJARCCE.2016.54278

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

III. LITERATURE REVIEW

The study in [4] provided an analysis of the critical factors affecting the energy consumption of mobile clients in cloud computing. Further, they have presented measurements about the central characteristics of contemporary mobile handheld devices that define the basic balance between local and remote computing. They also described a concrete example, which demonstrates energy savings. They also showed that the trade-offs are highly sensitive to the exact characteristics of the workload, data communication patterns and technologies used, and discuss the implications for the design and engineering of energy efficient mobile cloud computing solutions.

In [5] proposed a novel strategy for element union of VMs in view of versatile usage edges, which guarantees an abnormal state of meeting the Service Level Agreements (SLA). They have likewise accepted the high effectiveness of the proposed method crosswise over various types of workloads utilizing workload follows from more than a thousand Planet Lab servers. In [6] presented a multi-time period optimization model for saving the operational cost by combining two factors: 1) Dynamic Voltage/Frequency Scaling (DVFS), 2) turning servers on/off over a time horizon. They have also shown the impact of the granularity of the duration of the time slots and frequency options on optimal solutions. A parametric study on varying cost of turning servers on/off and power consumption is also presented.

In [7] considered the two constraints: deadline and budget. They proposed Deadline and Budget distribution-based Cost-Time Optimization (DBD-CTO) workflow scheduling algorithm that minimizes execution cost while meeting timeframe for delivering results and analyze the behavior of the algorithm. In case of Cloud computing, issues such as resource management and scheduling based on users' QoS constraints are yet to be addressed especially in the context of workflow management systems. Here author proposed a model for service placement in federated clouds to maximize profit while protecting quality of services (QoS) as specified in the service level agreements (SLA) of the workloads. Their contributions include an integer linear program (ILP) formulation of the generalized federated placement problem and application of this problem to load balancing and consolidation within a cloud.

In [8] dealt with this problem, presenting HCOC: The Hybrid Cloud Optimized Cost scheduling algorithm. HCOC decides which resources should be leased from the public cloud and aggregated to the private cloud to provide enough processing power to execute a workflow within a given execution time. They also presented extensive experimental and simulation results which show that HCOC can reduce costs while achieving the established desired execution time. An ideal cloud asset provisioning (OCRP) calculation is proposed in [9] to figure a stochastic programming model. The OCRP calculation can procurement registering assets for being utilized as a part of different provisioning stages and in addition a long haul arrangement, e.g., four stages in a quarter arrangement and twelve stages in a yearly arrangement. The interest and value instability is considered in OCRP. In this paper, distinctive methodologies are measured including deterministic proportionate plan, test normal guess, and Benders deterioration in OCRP calculation.

In [10], concentrated on the virtual machine position issue with an objective of minimizing the aggregate vitality utilization. A multi-dimensional space parcel model and a virtual machine position calculation were introduced. At the point when another VM situation assignment arrived, their calculation checked the back asset utilization state for each plausible PM, and after that picked the most suitable PM as per their proposed model to lessen the quantity of running PMs. In [12], author considered the issue of giving force planning support while managing numerous issues that emerged when spending plans virtualized frameworks. They oversaw power from a VM-driven perspective, where the objective was to know about worldwide utility tradeoffs between various virtual machines (and their applications) while keeping up power imperatives for the physical equipment on which they ran. Their way to deal with VM-mindful force planning utilized different circulated administrators coordinated into the virtual force administration (VPM) system.

In [13], introduced an essential hypothetical model and utilized it as a part of building overseeing, miniaturized scale lattices, and datacenter vitality administration. They investigated these divergent vitality administration frameworks and characterized a model for asset distribution that could be utilized for these and other vitality administration frameworks. The Datacenter Energy Management task was centered on demonstrating vitality utilization in server farms, with an objective to advance power utilization. Their task was centered on gathering information to characterize essential fuel utilization bends. In [14], tended to the issue of trimming so as to augment the incomes of cloud suppliers down their power costs. Approaches depended on element evaluations of client interest, and framework conduct models. A few approximations were utilized to handle the subsequent models. They had shown that choices, for example, what number of servers were controlled on can significantly affect the income earned by the supplier. Be that as it may, no startup power draw or execution assurances was considered.

In [15], exhibited Harmony, a Heterogeneity-Aware Resource Monitoring and administration framework that was equipped for performing dynamic limit provisioning (DCP) in heterogeneous server farms. Utilizing standard K means grouping, they demonstrated that the heterogeneous workload could be separated into numerous assignment classes with comparative qualities as far as asset and execution targets. The DCP was detailed as an advancement issue that considered machine and workload heterogeneity and also reconfiguration costs.

A structure used to naturally oversee processing assets of cloud foundations was proposed in [16] to at the same time accomplish suitable QoS levels and to decrease the measure of vitality utilized for giving administrations.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

Guazzone, Anglano and Canonico demonstrated that by means of discrete-occasion framework (DES) reproduction, their answer could oversee physical assets of a server farm in such an approach to essentially lessen SLO infringement regarding a conventional methodology. The vitality proficiency of the foundation was characterized as the measure of vitality used to serve a solitary application demand. To the best of our insight, applying the Npolicy for streamlining the mode-exchanging control and at the same time accomplishing the base expense under an execution ensure has not been considered some time recently.

IV. PROPOSED METHOD

This project proposed a for taking care of obliged improvement issue and making costs exchange offs in frameworks with various force sparing strategies. Distributed computing is another worldview for conveying remote figuring assets through a system. Be that as it may, accomplishing a vitality productivity control and all the while fulfilling an execution ensure have gotten to be basic issues for cloud suppliers. In numerous holding up line frameworks, the part of server is played by mechanical/electronic gadget, for example, PC, beds, ATM, Traffic light, and so forth., which is liable to inadvertent holding up of clients, it might fathomed by the servers excursion because of clump criteria. Ke et al. [6] examined the control strategy of the N-Policy M/G/1 line with server excursions, startup and breakdowns, where entry shapes a Poisson and administration times are by and large circulated. Power funds in cloud frameworks have been widely examined on different viewpoints lately, e.g., on the virtual machine (VM) side by moving VMs, applying solidification or assignment calculations, and on the server farm base side through asset portions, vitality administrations, and so forth.
The system architecture highlighted in Fig.2 shows that proposed model works in two processes. Briefly speaking, 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.

At the point when there has no employment in a holding up line or no occupation is being handled, a server gets to be sit without moving and it stays until a consequent employment has arrived. By and large, a server works then again between a bustling mode and an unmoving mode for a framework with irregular occupation entries in a cloud situation. A bustling mode shows that occupations are handled by a server running in one or a greater amount of its VMs'; and an unmoving mode demonstrates that a server stays dynamic yet no employment is being prepared around then. To moderate or take out unmoving force squandered, three force sparing arrangements with various vitality proficient controls, choice procedures and working modes are exhibited. In the first place, we attempt to make a vitality proficient control in a framework with three working modes m={Busy, Idle, Sleep}, where a rest mode would be in charge of sparing force utilization. A server is permitted to stay in an unmoving mode for a brief timeframe when there has no employment in the framework, as opposed to switch suddenly into a rest mode immediately when the framework gets to be unfilled [21]. An unmoving mode is the main working mode that associates with a rest mode. A server doesn't end its rest mode regardless of the fact that an occupation has arrived; it starts to work just when the quantity of employments in a line is more than the controlled N esteem. As per the exchanging process (from Idle to Sleep) and the vitality effective control (N arrangement), we have called such a methodology the "ISN approach".

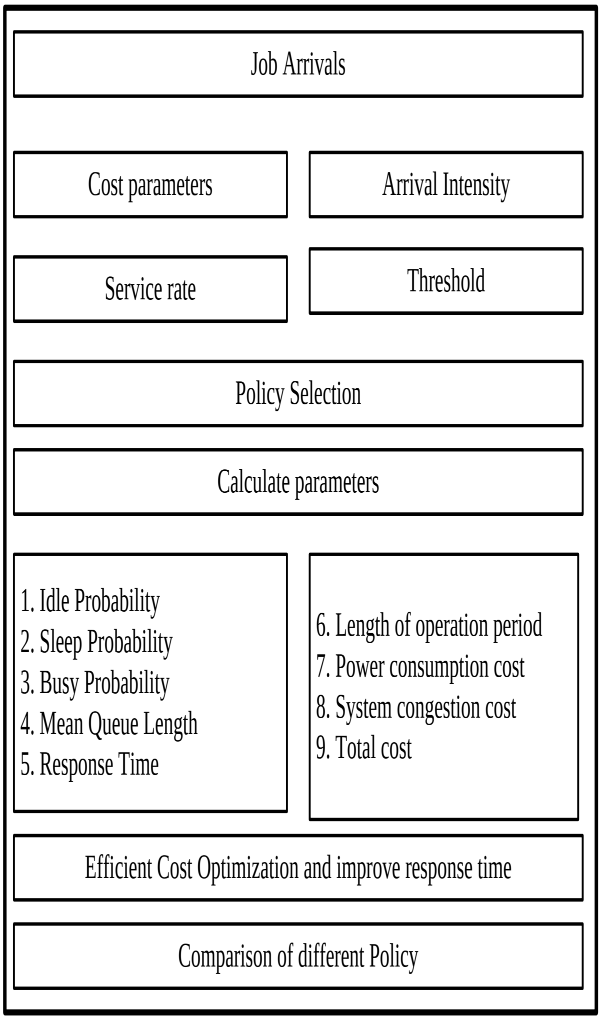

Fig.2 System Architecture of the Proposed Model

Fig.2 shows the regulated choice procedures and employment streams of the ISN strategy.

Step 1. A server closures its bustling mode when all present place of employment solicitations have been done.
Step 2. A server stays in an unmoving mode and sits tight for in this way arriving employments before exchanging into a rest mode.
Step 3. In the event that an occupation touches base amid an unmoving period, a server can switch into a bustling mode and begin to work instantly. A server starts a next

 DOI 10.17148/IJARCCE.2016.54278

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

unmoving period until all occupation demands have been effectively finished.

Step 4. In the event that there has no occupation entry, a server switches into a rest mode when an unmoving period lapses.

Step 5. A server stays in a rest mode if the quantity of occupations in the line is less than the controlled N esteem. Something else, a server switches into a bustling mode and starts to work.

Fundamentally, there have two instances of beginning a bustling mode:

Case 1. Beginning a bustling mode when a vocation lands in an unmoving mode;

Case 2. Beginning a bustling mode if the quantity of occupations in a holding up line is more than the N esteem when a rest period lapses.

Despite the fact that power is squandered in permitting a server to stay in the unmoving mode amid a non-load period, the advantages are that an entry work has more potential outcomes to get instantly administration and the server startup expense can be lessened.

V. RESULTS AND DISCUSSION

This session exhibits the numerical results to show the vitality proficiency of the proposed calculation. Here, we exhibited a vitality productive calculation for tackling obliged enhancement issues and making costs/exhibitions tradeoffs in frameworks with various force sparing approaches. The reproduction results are exhibited in this segment, it demonstrates that the advantages of lessening operational expenses and enhancing reaction times for various force sparing arrangements. This is executed utilizing Matlab tools. The result are show in figure from fig.3 to fig.7. For a perfect server in a server farm, power waste is intensified by the server itself, as well as the force conveyance misfortunes and cooling influence utilization, which builds influence utilization prerequisites by 50 to 100 percent. Subsequently, an unmoving cost parameter is settled inside of the reference range and the cost lattice is thought to be [C0 C1 C2 C3 C4 C5 C6] = [500, 400, 1, 8, 2, 3, 3] in recreations.

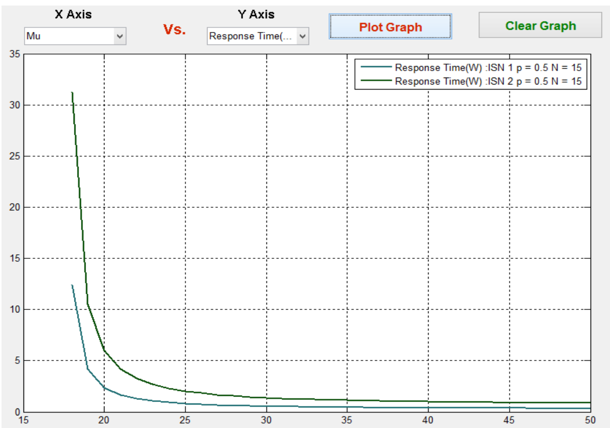

Fig 3 Plot of Response Time for ISN-1 and ISN-2 at N=15, ρ=0.5

As we can see, decreases out of gear likelihood between frameworks with various force sparing arrangements are not the same. A large portion of their switching so as to unmoving times can be diminished or eliminated into rest modes for frameworks with the force sparing strategies as appeared in fig.6.3 for both polices with N=15, ρ=0.5.

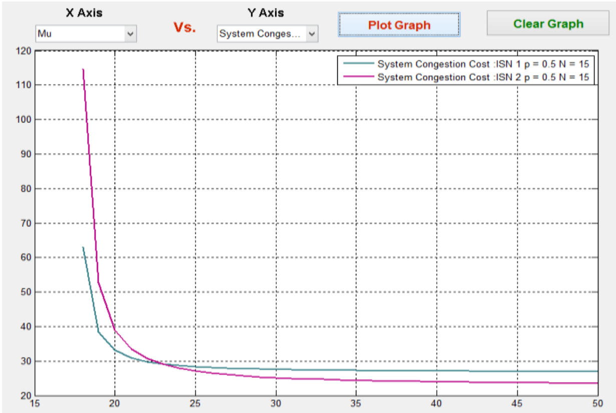

Fig 4. Plot System Congestion Cost for ISN-1 and ISN-2 at N=15, ρ=0.5

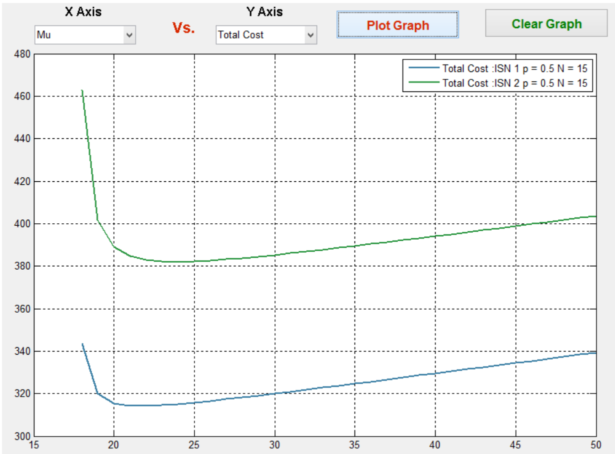

Fig 5. Plot Total Cost for ISN-1 and ISN-2 at N=15, ρ=0.5

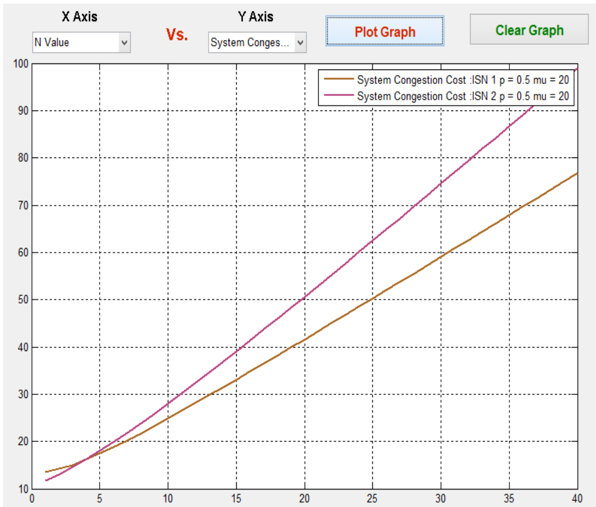

Fig 6. Plot System Congestion cost for ISN-1 and ISN-2 at mu=20, ρ=0.5

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 5, Issue 4, April 2016

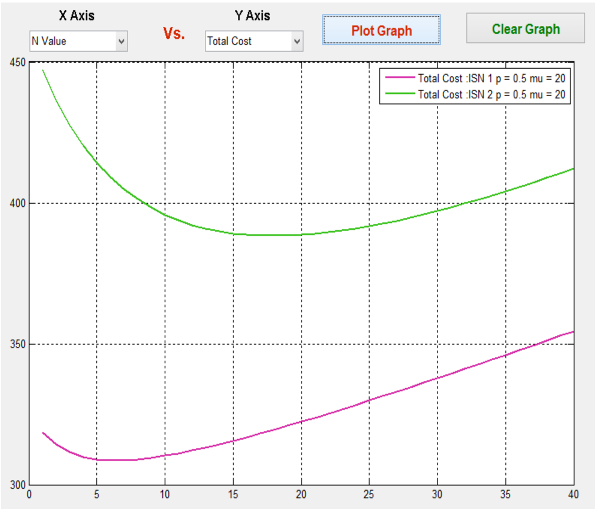

Figure 7. Plot Total Cost for ISN-1 and ISN-2 at mu=20, ρ=0.5

VI. CONCLUSION AND FUTURE DIRECTION

The growing crisis in power shortages has brought a concern in existing and future cloud system designs. To mitigate unnecessary idle power consumption, three power saving policies with different decision processes and mode switching controls are considered. 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. The issue of choosing a suitable policy among diverse power managements 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 significantly improve the response time in a low arrival rate situation. On the other hand, applying others policies can obtain more cost benefits when the startup cost is high. As compared to a general policy, cost savings and response time improvement can be verified.

REFERENCES

[1] A. P. Azad, S. Alouf, E. Altman, V. Borkar and G. Paschos, "Vacation policy optimization with application to IEEE 802.16e power saving mechanism," Wireless Days (WD), 2009 2nd IFIP, Paris, 2009, pp. 1-7.

[2] Y. J. Chiang, Y. C. Ouyang and C. H. (. Hsu, "An Efficient Green Control Algorithm in Cloud Computing for Cost Optimization," in IEEE Transactions on Cloud Computing, vol. 3, no. 2, pp. 145-155, April-June 1 2015.

[3] L. Kong and D. H. K. Tsang, "Optimal Selection of Power Saving Classes in IEEE 802.16e," Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE, Kowloon, 2007, pp. 1836-1841.

[4] G. Zangara, D. Terrana, P. P. Corso, M. Ughetti and G. Montalbano, "A Cloud Federation Architecture," 2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), Krakow, 2015, pp. 498-503.

[5] A. Dalvandi, M. Gurusamy and K. C. Chua, "Time-Aware VMFlow Placement, Routing, and Migration for Power Efficiency in Data Centers," in IEEE Transactions on Network and Service Management, vol. 12, no. 3, pp. 349-362, Sept. 2015.

[6] F. C. Jiang, Der-Chen Huang, Chao-Tung Yang and Kuo-Hsiung Wang, "Mitigation techniques for the energy hole problem in sensor networks using N-policy M/G/1 queuing models," Frontier Computing. Theory, Technologies and Applications, 2010 IET International Conference on, Taichung, 2010, pp. 281-286.

[7] F. C. Jiang, C. T. Yang, K. H. Wang and D. C. Huang, "Design Framework to Optimize Power Consumption and Latency Delay for Sensor Nodes Using Min(N, T) Policy M/G/1 Queuing Models," Future Information Technology (FutureTech), 2010 5th International Conference on, Busan, 2010, pp. 1-8.

[8] E. Pakbaznia, M. Ghasemazar and M. Pedram, "Temperature-aware dynamic resource provisioning in a power-optimized datacenter," Design, Automation & Test in Europe Conference & Exhibition (DATE), 2010, Dresden, 2010, pp. 124-129.

[9] Huan Chen and Chen-Wei Huang, "Power management modeling and optimal policy for IEEE 802.11 WLAN systems," Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th, 2004, pp. 4416-4421 Vol. 6.

[10] I. Kamitsos, P. Tsiaflakis, K. Kerpez, S. Ha and M. Chiang, "Energy efficient DSL via heterogeneous sleeping states: Optimization structures and operation guidelines," Global Communications Conference (GLOBECOM), 2012 IEEE, Anaheim, CA, 2012, pp. 3128-3134.

[11] V. Eramo; E. Miucci; M. Ammar, "Study of Reconfiguration Cost and Energy Aware VNE Policies in Cycle-Stationary Traffic Scenarios," in IEEE Journal on Selected Areas in Communications , vol.PP, no.99, pp.1-1.

[12] F. Angiolini, L. Benini and A. Caprara, "An efficient profile-based algorithm for scratchpad memory partitioning," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 24, no. 11, pp. 1660-1676, Nov. 2005.

[13] V. Eramo, E. Miucci, L. Compagnone and M. Ammar, "Evaluation of reconfiguration cost and Energy Aware Virtual Network Embedding policies," Communication Workshop (ICCW), 2015 IEEE International Conference on, London, 2015, pp. 2824-2829.

[14] S. O. D. Luiz, A. Perkusich, B. M. J. Cruz, B. H. M. Neves and G. M. da S. Araujo, "Optimization of timeout-based power management policies for network interfaces," in IEEE Transactions on Consumer Electronics, vol. 59, no. 1, pp. 101-106, February 2013.

[15] Bu Aiguo, Shi Longxing, Hu Chen, Li Jie and Wang Chao, "Energy-optimal dynamic voltage scaling for sporadic tasks," Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on, Island of Kos, 2006, pp. 4.

[16] T. Horvath, K. Skadron and T. Abdelzaher, "Enhancing Energy Efficiency in Multi-tier Web Server Clusters via Prioritization," Parallel and Distributed Processing Symposium, 2007. IPDPS 2007. IEEE International, Long Beach, CA, 2007, pp. 1-6.

BIOGRAPHIES

Banashankari received the BE in Information Science and Engineering at SIET, Bijapur. She is final year M.Tech student in Computer Science and Engineering at EWIT, Bengaluru. Her current area of research is in Cloud Computing.

Chandan Raj B R is working as Assistant Professor, Department of Computer Science, East West Institute of Technology, Bengaluru-91. His areas of specialization are Mobile Computing, 4G Network Management, Sensor Networks and Network Security