Energy Calculation in a Data Center with the Losses and the Method Proposed to Reduce Energy Consumption

Nazneen Taj
Asst. Prof.,
Dept. Of CSE KNSIT,
Bangalore-560064

Dr. Anirban Basu
Prof. Dept. Of CSE APSCE,
Bangalore, India.

Abstract:- **Today, energy conservation is gaining lot of attention. The subject Green IT, Sustainability has emerged which involves the study and practice of environmentally sustainable computing and includes "designing, manufacturing, using, and disposing of computers, servers, and associated subsystems-such as monitors, printers, storage devices, and networking and communications systems efficiently and effectively with minimal or no impact on the environment". When we calculate the energy utilization of a data center we have to consider both energy consumption in server systems and the energy required for cooling and in power delivery infrastructure. We need to calculate the energy consumption by each server by using basic power consumption equation so that we can calculate the total energy consumed by the data center. Some authors have calculated the energy consumption by a processor during idle time and during busy time while some have proposed scheduling algorithms which involve varying the voltage and frequency based on requirements. Algorithms like DVFS can schedule voltage and frequencies of CPUs in low power active state and suspended inactive state. We have focused on calculating not only the losses that occur during scheduling but also the losses caused by the components used. Authors have proposed equations to calculate the total energy consumed without considering the losses. We have discussed how energy supplied at different times or different intervals of time can be saved. Some measures have been proposed to minimize the losses.**

Keywords: - Cloud computing, virtual machine, task scheduling, genetic algorithm, group search optimization algorithm, GGSO, Infrastructure as a Service (Iaas), Quality –of-Service (QoS).

1. INTRODUCTION

A data center where servers consume enormous amount of energy and release carbon dioxide poses serious environmental hazards. The energy consumption in a data center with large pool of servers determines the maximum size and computing capacity of the data center and therefore the maximum service that can be provided to clients. Studies reveal that data centers of the IT industry contribute nearly 2% out of 14% of total carbon emission from industries. This is creating serious environmental hazards including global warming. With the advent of Cloud Computing using data centers with a large number of servers, the problem is becoming more serious and there is a need for efficient energy management.

Cloud Computing has as of late developed as an appealing model of offering Information Technology (IT) foundation (i.e., registering, stockpiling, and system) to expansive and in addition little endeavors both in private and open areas. Cloud administration suppliers offer these administrations focused around altered Service Level Agreements (SLAs) which characterize client's obliged Quality of Service (Qos) parameters. Distributed computing decreases speculation in different assets like equipment, programming and permit assets to be rented and discharged. It decreases beginning speculation, support cost and working expenses. Cloud administrations are facilitated on administration supplier's own particular framework or on outsider cloud foundation suppliers. As a critical piece of distributed computing, Infrastructure as a Service (Iaas) gets to be exceptionally well known as the establishment for larger amount administrations. Chiefly, three sorts of administrations are conveyed: Platform as a Service (Paas), Infrastructure as a Service (Iaas) and Software as a Service (Saas). Cloud clients utilize these administrations at whatever point required by interest utilizing pay-every use model. Iaas suppliers, for example, Amazon EC2 and IBM Smart Cloud Enterprises, permit clients to lease assets as Virtual Machines (VMs). They can offer distinctive VM sorts that are described by machine arrangement, Qos and estimating model. One average agent is Amazon EC2 which can give three sorts of evaluating models: on-interest, reticent and spot. objectives of this research include:

1. In-depth study on energy utilization and its importance in data centers.
2. Study on existing energy evaluation algorithms such as Dynamic Voltage Frequency Scaling(DVFS), Group search optimization(GSO) and Genetic Algorithm.

3. Developing methodology to enhance DVFS algorithm by considering losses of its equipment along with voltage and frequency in the data center.
4. Deriving improved version of energy equations for optimum utilization of energy
5. Development of hybrid model using GA and GSO for effective utilization of energy
6. Finding an algorithm which works on dynamic load, increases profit of the cloud providers and saves appreciable amount of energy.

2. LITERATURE REVIEW

Good sleep scheduling and virtualization techniques of computing resources in data centers improves the energy efficiency of data center. The cloud, in essence, comprises a group of virtual machines equipped with the facility of computation and storage. We are working on a task scheduling scheme on diverse computing systems using a hybridization of genetic and Group search optimization algorithm (GGSO) is proposed. The basic idea of our approach is to exploit the advantages of both genetic algorithm (GA) and group search optimization algorithms (GSO) while avoiding their drawbacks. In GGSO, each dimension of a solution symbolizes a task and a solution as a whole signifies all tasks priorities. The important issue is how to assign users tasks to maximize the income of Infrastructure as a Service (Iaas) provider while promising Quality-of-Service (QoS). The generated solution is competent to assure user-level (QoS) and improve IaaS providers' credibility and economic benefit. The GGSO method also designs the producer, scrounger ranger, crossover operator and suitable fitness function of the corresponding task. According to the evolved results, it has been found that our algorithm always out- perform the traditional algorithms. Cloud Computing has as of late developed as an appealing model of offering Information Technology (IT) foundation (i.e., registering, stockpiling, and system) to expansive and in addition little endeavours both in private and open areas [1,2]. Cloud administration suppliers offer these administrations focused around altered Service Level Agreements (SLAs) which characterize client's obliged Quality of Service (Qos) parameters. Distributed computing decreases speculation in different assets like equipment, programming [3] and permit assets to be rented and discharged. It decreases beginning speculation, support cost and working expenses. Cloud administrations are facilitated on administration supplier's own particular framework or on outsider cloud foundation suppliers [4]. As a critical piece of distributed computing, Infrastructure as a Service (Iaas) gets to be exceptionally well known as the establishment for larger amount administrations [5]. Chiefly, three sorts of administrations are conveyed: Platform as a Service (Paas), Infrastructure as a Service (Iaas) and Software as a Service (Saas). Cloud clients utilize these administrations at whatever point required by interest utilizing pay every use model. Iaas suppliers, for example, Amazon EC2 and IBM Smart Cloud Enterprise [6], permit clients to lease assets as Virtual Machines (VMs). They can offer distinctive VM sorts that are described by machine arrangement, Qos and estimating model. One average agent is Amazon EC2 which can give three sorts of evaluating models: on-interest, reticent and spot [7, 8]. the biggest concerns about cloud computing are security and privacy . The traditional security problems such as security vulnerabilities, virus and hack attack can also make threats to the cloud system and can lead more serious results because of property of cloud computing. Hackers and malicious intruder may hack into cloud accounts and steal sensitive data stored in cloud systems. The data and business application are stored in the cloud center and the cloud system must protect the resource carefully [21]. Moreover, the clouds can be classified in to three categories such as public, private and hybrid [9]. At the point when the cloud becomes accessible for the common client on a pay-every use stipulation, it is treated as the public cloud. At the point when associations build their own specific applications and operate their own specifics in the structure it is labelled as the private cloud as right of entry is limited to clients within the association. The hybrid cloud is furnished by the integration and amalgamation of the public and private clouds [10]. A power model has been proposed by authors [11] which is given as,

$E=\int_t P.U(t)$ ……………………….(1)

Where, P is the power utilized and U(t) is the usage of the machine with time. As time increases energy consumption also increases and we need to calculate the losses which increase as the usage increases. Energy calculations have been done by authors in [12] incorporating busy CPU and idle CPU, times and proposed overall energy consumption equations like

Eprocessor=busy time *Pbusy+ (system time – busy time) * Pidle --------- (2)

The overall energy consumption of the center with n processors can be estimated as proposed by authors

in[13]
$E_{overall}=\Sigma_{n=1}E_{processor}$ ……………………….(3)

Authors [14] have calculated the energy consumed by a data center as the integral power consumed with respect to time but authors[11] have shown Pdep = CefV ^{2} dd

From Equation (1) E=

If we consider the losses, then the equation (1) is modified as

$$E=\int_{t1}^{t2} PU(t) \text{ *losses} \quad \ldots\ldots(4)$$

Where, E is the energy utilized which is equal to the overall power utilized with respect to time multiplied by the heat losses. Assuming as the load of a data center increases with time the amount of power consumed by a data center also increases. The power consumption of a data center also increases the power requirement. After Calculating, the accurate amount of energy utilized in the data center we worked and studied on some algorithms which were based on task scheduling and minimizing the energy consumption. We studied the Genetic algorithm and group search optimization algorithm[13][14].

A. Genetic algorithm

A genetic algorithm was first launched by Holland in 1975[15], as an iterative stochastic in which the natural appraisal is employed to mould the search technique. The GA techniques can be employed to tackle the optimization issues by replicating the genetic function of the biological organism [16, 17]. As indicated by the name, the GA approaches follow the evolutionary theory in nature to tackle the optimization issues. A general genetic algorithm performs three genetic functions such as the *selection*, *cross over* and *mutation*. In the selection, certain solutions from the populations are chosen as parents where as in the crossover the parents are crossbred to generate the offspring. In the case of the mutation the offspring is organized in accordance with the mutation rules. In the genetic algorithm, a solution is termed as the individual and the iterations of the algorithm are labelled as the generation.

B. Group search optimization algorithm

The GSO technique is at first proposed by [9]. It is a population –based optimization approach and uses the producer-scrounger (PS) model and animal scanning mechanism. Producer-scrounger model for designing optimum searching strategies was stimulated by animal searching conduct and group living theory [18]. It has often paved the way for the adoption of two foraging techniques within groups: 1) creating and looking for food; and 2) scrounging, joining resources exposed by others. With a view not to entrap in local minima, the GSO further deploys the “ranger” foraging techniques.

The population of the GSO approach is known as a group and each individual in the population is termed a member.

3. PROPOSED WORK

We have proposed an equation to calculate the losses and energy consumed. This equation helps to know how much power we need to supply to the data center and what are the other losses that occur excluding the calculated one. As we know that the power required by a data center is not constant but depends upon the load and duration of usage of servers or systems[19]. When we want to know how much energy is to be supplied to a data center we need to know the heat generated and the requirement of cooling systems. When we calculate the total energy requirement in a data center we need to calculate the losses that occur in a data center. We propose the following equation:

$$ET=Ec+TE+\int t \text{ PCPU}+HL \ldots\ldots (5)$$

Where, ET stands for total energy consumed by data center, EC is the energy required for cooling, TE is the minimum energy consumed by both idle and active CPUs in data center, ∫tPCPU is the power consumed by the CPU with respect to time. HL is the heat loss caused by the components used in the data center. EC can be calculated as EC=NS*HG in joules

Where, NS stands for minimum supply given to cooling system and HG are the heat losses.

By considering the redundancy, the factor which is considered for the failure or servicing of the cooling equipment of any other equipment in a data center [20].

Therefore if we consider having redundant equipment (equipment which considers redundancy) we can re-write the energy equation as

$$EC=NS+HG+R \ldots\ldots(6)$$

Where, R stands for redundancy. If we have 100KW of power usage then we need 100KW of cooling plus redundancy to allow for cooling equipment failures and servicing. We may also need to add additional cooling capacity to allow electronic equipment lay out or computer room infrastructure design like load cooling delivery or overloaded areas.

TE is the minimum energy consumed by both idle and active CPUs in data center which is given by TE= PI+PA where PI is the power consumed by idle CPU and PA is the power consumed by active CPU. HL is the heat loss caused by the components used in the data center given by.

$$HL= \int t \text{ LD. VD}+fD \ldots\ldots (7)$$

Where LD is the load on the data center, VD and fD are variations in voltage and frequency respectively.

VD is given by VD=V max -Vmin, which are maximum and minimum voltages respectively.

fD is given by f D=f max-f min, which are maximum and minimum frequencies respectively. We have to consider the losses which are function of time i.e. longer the CPU runs more is the power required. Therefore longer the CPU runs more will be the heat generated from CPU which can be calculated by

$$\int t \text{ } P_{CPU} \text{*}HL \ldots\ldots(8)$$

The above equations are implemented based on the values taken from the referenced authors papers [18][14][4] and we have got the following graphs.

4. RESULTS

The results plotted as fig.1 to fig. 3 shows that when we use our proposed equation to calculate the amount of energy consumed in a data center it is giving the better results when compared to the old equations. The results also shows that if we know the exact amount of energy consumed by the data center by minimizing the losses the we don't need to run CPUs for longer hours and thus can save good amount of energy. When we know that what is actual amount of energy required by the data center and avoid the losses we can find that our equation gives the less numbers of hours the CPU should be running to complete the same task.

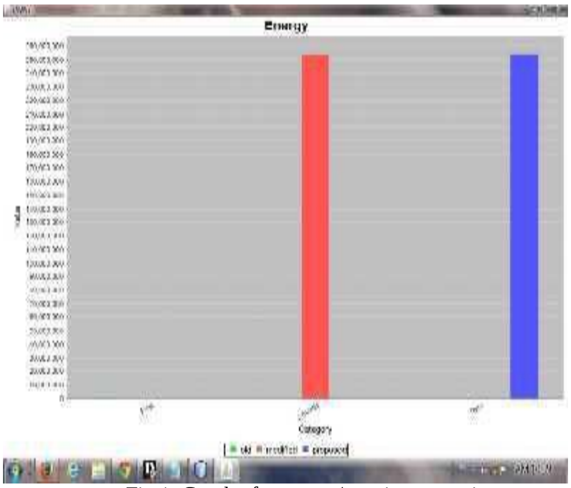

Fig.1: Graph of energy v/s various equations.

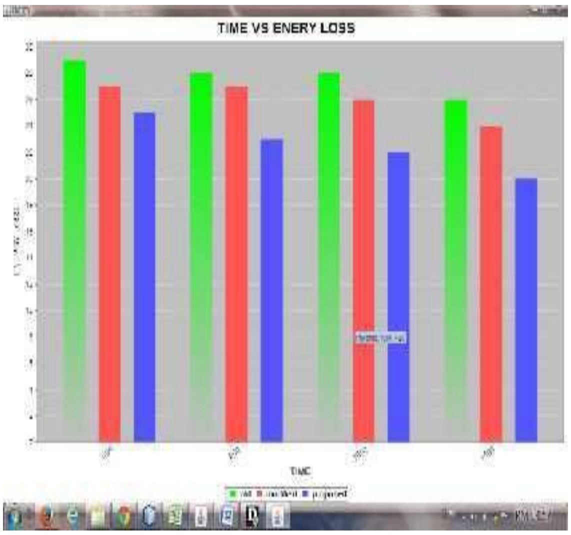

Fig.2: Graph of CPU v/s time duration of energy consumption

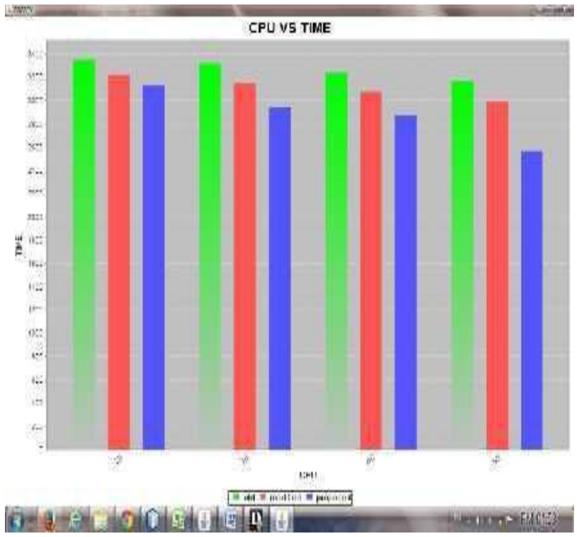

Fig.3: Graph of Time v/s energy loss.

Thus we have proposed an equation which calculate the amount of energy consumed by the data center by considering the losses in the data center. We have taken some values of energy consumed and losses from the papers of the authors who have worked on the same and we found that our equations calculate much accurate energy consumed values. After proposing an equation we are trying to find an algorithm to reduce the energy consumption in a data center. We are working on Genetic algorithm and Group search optimization algorithm and trying to come up with a hybrid technique called GGSO.

5. PROPOSED ALGORITHM GGSO (Genetic Group Search Optimization) Algorithm

Based Task Scheduling

The objective of this research is to schedule the task based on the GGSO algorithm. Here, our hybrid optimization algorithm is designed between the genetic algorithm [19] and group search optimization algorithm [20]. Our approach is exploiting the advantages of the genetic algorithm and group search optimization algorithm while avoiding their draw backs. By hybridized these two optimization algorithm, it will be overcome the disadvantages of the individual performance of the GSO and GA and it has the advantages of easily realizing and quickly converging, so that this scheduling approach is able to get an optimal or suboptimal solution in a shorter computational time. Hybridizing

GSO and GA leads to the combined effect of the good global search and good local search algorithm, which yields a promising result. In our approach at first, we apply the GSO algorithm to the application, which find out the optimal task scheduling to the corresponding application. Finally in each iteration, the worst solution is replaced by crossover operator to improve the solution quality. We are working on the implementation of the Algorithm for GGSO in task scheduling where we are planning these following algorithmic steps to be implemented.

GGSO Algorithm

Input

Parameter for the GGSO algorithm; Parameter for the task scheduling; Output
A task scheduled

1.call step 1 to create an initial solution
2. repeat
3. call algorithm 2 to evaluate the fitness function
4. copy the elitism directly to the next new solution
5. repeat
6. call step 3 to select a producer for the task
7. call step 4 to select the scrounger of the task
8. call step 5 to select the rangers
7. call algorithm 4 to replace a worst solution by crossover operator
8. until the new population complete
9. replace the old population by the new population
10. until the termination condition is satisfied
11. return optimal scheduled task.

Our final research work will be the implementation of Hybrid algorithm GGSO for task scheduling and thus to save appreciable amount of energy consumption.

6. CONCLUSION

Energy utilization is one of the challenging research issue where various works have been reported in the existing literature. Effective utilization of power play major role in various application especially in data centers. This research work focuses on development of an effective model for utilizing energy in data centers. A new energy equation has been proposed in this work. A GSO and GA based hybrid model has been proposed in this work for effective energy consumption. Results show increased power utilization using proposed model.

7. FUTURE WORK

- Energy equations proposed never considered the losses that occurred in a data center.
- Those equations does not hold good for the variable loads.
- Genetic algorithm converges slowly when it is used individually.
- GSO algorithm may loose some solutions while finding the best path to reach the optimal solution.

ACKNOWLEDGEMENT

I would like to tribute my success, achievement and hard work to my family members, my guide, mentors, friends and also the people who criticized me. Only because of them I could understand where and what my weaknesses are , I'm thankful to all of them and try hard and hard to go further. I would like to thank the conference organizers and all the journal editors who make us to do the quality of work.

REFERENCES

[1] M. Armbrust, et al., "A View of Cloud Computing," Communication of ACM, vol. 53, pp. 50-58, 2010.
[2] 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 5th Utility", Journal of Future Generation Computer Systems, Vol.25, No. 6, pp. 599–616, 2009.
[3] J. Geelan, "Twenty one experts define Cloud computing Virtualization", Electronic magazine, [Online] Available at http:// virtualization.sys.con.com/node/612375.
[4] L. Mei, W.K. Chan and T.H. Tse, "A Tale of Clouds: Paradigm Comparisons and some Thoughts on Research Issues", IEEE Asia-Pacific Services Computing Conference (APSCC), pp. 464-469, 2008.
[5] S. Bhardwaj, L. Jain, and S. Jain, "Cloud computing: A study of infrastructure as a service (IaaS)," Int. J. Eng. Inf. Technol., vol. 2, no. 1, pp. 60–63, 2010.
[6] R. Buyya, C. S. Yeo, S. Venugopal, and J. Broberg, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," Future Generation Comput. Syst., vol. 25, no. 6, pp. 599–616, 2009.
[7] A. N. Toosi, R. N. Calheiros, P. K. Thulasiram, and R. Buyya, "Resource provisioning policies to increase IaaS provider's profit in a federated cloud environment," in Proc. IEEE Int. Conf. High Perform. Comput. Commun., Banff, Canada, 2011, pp. 279–287.
[8] H. Liu and D. Orban, "GridBatch: Cloud computing for large-scale data-intensive batch applications" , in Proc. IEEE Int. Symp. Cluster Comput. Grid, pp. 295–305, 2008.
[9] S. He, Q. H. Wu and J. R. Saunders, "A Group Search Optimizer for Neural Network Training" Lecture Notes in Computer Science, vol.3982, pp.934-943, 2006.
[10] LI Wenhao, "A Community Cloud Oriented Workflow System Framework and its Scheduling Strategy", IEEE 2nd Symposium on Web Society (SWS), 2010.
[11] Hang Yuan, Ishfaq Ahmad, "Energy efficiency in data center and cloud based multimedia services : An overview and future directions, IEEE,978-4244-7614-5/10, 2009.
[12] Mohammed A. Haque, Dakai Zhu, "Energy aware standby technique for periodic real time applications ", IEEE, 978-1-4577-7/11,2011.
[13] Tomoaki Hamano, Toshio Endo and Satoshi Matsuoka, "Power aware dynamic task scheduling for heterogeneous accelerated clusters", IEEE, 978-1-4244-3750-4/09, 2009.
[14] RajkumarBhuyya, Anton Beloglazov and JemalAbawajy," Energy efficient management of data center resources for cloud computing: A vision ,Architecture elements, and open challenges, IEEE, 2011.
[15] J. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, MI, USA, 1975.
[16] S. Song, K. Hwang and Y.-K., "Kwok, Risk-resilient heuristics and genetic algorithms for securityassured grid job scheduling", IEEE Trans. Computing, Vol.55, no.6, pp.703–719.
[17] A. Zomaya, C. Ward, B. Macey, Genetic scheduling for parallel processor systems: comparative studies and performance issues, IEEE Trans. Parallel Distributing System, vol. 10, no.8, pp.795–812, 1999.
[18] I. D. Couzin, J. Krause, N. R. Franks, and S. A. Levin, "Effective leadership and decision-making in animal groups on the move," Nature, vol. 434, pp. 513–516, 2005.
[19] Fatma A. Omara, Mona M. Arafa,"Genetic algorithms for task scheduling problem," J. Parallel Distrib.Comput. Vol.70 ,pp.13-22, 2010.
[20] S. He, Q. H. Wu and J. R. Saunders, "Group Search Optimizer: An Optimization Algorithm Inspired by Animal Searching Behavior", IEEE transactions on evolutionary computation, vol. 13, no. 5, 2009.
[21] Raj Kumar,"Research on cloud computing security threats using data transmission",IJARCSSE,vol. 5,issue 1,Jan 2015.

AUTHORS BIOGRAPHY:

Nazneen Taj, is doing Ph.D in Bharathiar university, Registered in 2011 and expecting her degree in 2017. She did her M-Tech in computer science & Engg. In UBDTCE, Davanagere, Karnataka & did her B.E. in SJMIT, Chitradura., Karnataka. She is severing education field from 12 years. She served in SJMIT, Citradurga, Karnataka from 2004 to 2010. She joined CMRIT, Bangalore, Karnataka and served from 2010 to 2011. She is working in KNSIT, Bangalore, Karnataka from 2011 to till date. She is also working as Placement and Training officer in KNSIT, and helping the student community to get placed. She is working as a consultant to a company M/S OOPs Labs. Her area of interest is Cloud computing, Electronic circuits, Network analysis, software testing, computer networks, Microprocessor, Embedded systems, Logic Design, Artificial Intelligence etc.

Dr Anirban Basu has been elected as the President of Computer Society of India (CSI) for the period 2016-2017. He will serve as the Vice President of CSI from April 2015 to 2016 before assuming the office of the President. Formed in 1965, CSI is the largest and oldest body of IT professionals and has been instrumental in guiding the Indian IT industry down the right path since its formative years. Today, CSI has 73 chapters all over India, over 500 student branches, and more than 100000 members including India’s most famous IT industry leaders, brilliant scientists and dedicated academicians. Dr Anirban Basu is a dynamic executive based in Bangalore with more than 35 years of experience in Academia, advanced Research & Development, commercial Software Industry, Consultancy and Corporate Training.