Indian Journal of Science and Technology, *Vol 9(S1), DOI: 10.17485/ijst/2016/v9iS1/103143, December 2016*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Optimizing the Energy Efficiency of the Modern Data Centers using Ant Colony Optimization

Jaspreet Singh Gill* and Geet Kiran Kaur

Chandigarh University, Gharuan - 140413, Punjab, India; j.s.gill@outlook.com, GeetKiran.cse@cumail.in

Abstract

Objectives: To propose an energy efficient Swarm based Optimization method. **Methods/Statistical Analysis**: In this research, a set of tasks and servers are taken as Input. The data center server energy consumption is taken as the output of the algorithm. The processing time may vary according to the number of tasks given. Task allocation is done in such a way that most-efficient-server gets the tasks first. If, average job density is low, i.e. then it works on the same system and if the value is high then it moves to next one. The algorithm follows a scheduling and divides the jobs into servers and start execution of the job tasks. **Findings**: In this paper, an ant colony algorithm is proposed to efficiently allocate tasks to virtual machines, which allocates resources based on the available resources and the energy consumption of each virtual machine. This ACO algorithm is implemented, executed, and evaluated using the experiments in Cloud Sim. Reduced power consumption, with throughput has been obtained. **Application/Improvements:** Less SLA violation has been obtained with good response time.

Keywords: Ant Colony Optimization, Cloud Sim, Energy Consumption, Green Cloud Computing, Optimization, Virtualization

1. Introduction

The rising usage of the cloud computing leads to the adoption of new features like online application, online data storage via web[1]. Cloud computing has also been adopted by various organizations like Amazon, Rackspace for better resource utilization. Also, the clients of these organizations are moving from physical servers to virtual servers. Energy consumption is the main key factor of these companies based on the data centers. A large number of requests at server side give rise to various traffic and energy consumption at the high rate. Also power as, well as energy usage by companies, constitutes to 1.4 % of total electrical energy in world[2,3].

There is data center named Barcelona and it devotes 1.2 MV of energy consumption[4] which is equal to 1200 houses. Minimization of energy usage is a very critical factor these days and must be resolved inaccurate way. Bigger data centers have a high consumption of energy. So, they must be optimized.

To manage the issue of energy consumption there is strong need to manage the QOS parameters. So the primary goal of this work is to minimize the energy consumption using adoption of SLA violation.

1.1 Energy Consumption Pattern

The issue of energy utilization in IT sector has been accepting expanding consideration lately and there is developing acknowledgment of the need to oversee energy utilization over the whole Information and Communications Technology (ICT) sector. In the last propel of years, Cloud[5] based server centers are expanding incredibly due to the interest for PC asset. Since more server centers are started to be the energy utilization, of these server centers are likewise expanded as it were. Notwithstanding high energy utilization there is an expansion effect on nature by the type of Carbon Dioxide discharges. As indicated by the report to Congress on Server and server centers[5], the server centers are in charge of around 2% of worldwide CO_2 outflow and they utilize almost 80 million megawatt-

**Author for correspondence*

hours of energy yearly, it is around 1.5 times the measure of power utilized by the entire New York City[6].

By 2020 the aggregate sum of Carbon Dioxide discharged by these server centers will be about 359 megatons. In such a circumstance it is significant significance that the cloud server centers ought to have decent energy productivity[7]. The Major issue in poor energy productivity is that a large portion of the energy is squandered when servers keep running at low usage.

As per the late research from Pike Research[8], the worldwide business for green server centers will develop from $17.1 billion in 2012 to $45.4 billion by 2016. Indeed, on location server with no virtualization will emanate session 46 kg of CO_2 every year.

1.2 Methods of Green Computing

1.2.1 Lower Power Hardware

Computers can be made less energy consumption devices by using lower power processor, using cooling devices as well as using spinning SSD of small size rather than large size[9]. Intel has developed a process that is going to use less power. Low power PCs are green saver so they don't allow fast gaming.

1.2.2 Virtualization

Virtualization is the procedure of makes utilization of effective framework assets. Virtualization fits in green computing by joining servers and boosting CPU throughput. With virtualization a few physical PCs can be made one virtual PC on single PC[10]. Points of interest of Virtualization:

- Efficient utilization of assets
- Superior level of deliberation
- Replication
- Scalable and adaptable framework

This empowers much of the time got to records to be put away on superior, low-limit drives; whilst documents in less utilize are set on more power-proficient, low-speed, bigger limit drives[11].

1.2.3 Internet based Applications

Green registering is not simply like taxpayer supported organizations that are urging to reconsider their utilization in IT segment. Be that as it may, it has ended up reality in numerous nations. In web based applications, it systems directors in decreasing CO_2 outflow[12].

1.2.4 Storage

There are three courses accessible for capacity of information that makes productive utilization of assets. So as to have less power utilization, there is requirement for advancement of power gadgets. So it can be accomplished by 3.5' hard drive ease of use or either 2.5" hard circle convenience. Most minimal power utilization is likewise accomplished by utilizing SSD cards.

1.2.5 Maximization of Virtual Machines

Virtualization, a strategy to run a few working frameworks at the same time on one physical server, has turned into a center idea in present day datacenters, predominantly determined by advantage of use seclusion, asset sharing, adaptation to internal failure, transportability and cost effectiveness. A unique middleware, hypervisor, abstracts from physical equipment assets and gives alleged virtual machines acting like genuine PCs with their own (virtual) equipment assets. Flexibility and asset usage is being accomplished by VM. Virtual Machines (VMs) can be powerfully allotted and de-dispensed to the assets as per the interest and accessibility, and the likelihood of merging various VMs into the same physical server. In any case, it might happen that the accessible assets are insufficient for the present interest at a given top time, particularly on account of private greens. In these cases, the green blasting strategy is utilized to get on-interest VMs (regularly) from general society green, in this manner developing the processing limit of the office straightforwardly to the client. Improvement of current designation of VMs is completed in two stages: at the initial step we select VMs that should be relocated, at the second step picked VMs are set on the host. We for the most part utilize two heuristics for picking VMs to move[13].

1. The first heuristic, Speedup, depends on quick handling while keeping to protect free assets with a specific end goal to avert SLA infringement because of solidification in situations when asset prerequisites by VMs increment. At every time span all VMs are reallocated. New position is accomplished by live movement of VMs.

2. The second heuristic i.e. Vitality depends on aggregate usage of CPU by all VMs between these Set

qualities. On the off chance that the use of CPU for a host falls beneath the lower esteem, all VMs must be moved from this host and the host must be exchanged off with a specific end goal to take out unmoving power utilization. In the event that the usage goes over the upper esteem, some VMs must be relocated from the host to lessen use to anticipate potential SLA infringement[14,15].

1.2.6 Wireless Sensor Network

A system is a gathering of two or more PC frameworks which are connected together to speak with each other. It is a telecom system that permits PCs to trade information. In PC systems, arranged registering gadgets pass information to each other along information associations. The associations between hubs are set up utilizing either link media or remote media. The best-known PC system is the Internet. Diverse frameworks offer assets accessible in the system. Shared assets can be of programming sort or equipment sort. The gadgets that structure system to trade information are called system hubs. These hubs can incorporate has, for example, PCs, telephones, servers and also organizing equipment.

PC systems vary on the premise of physical media used to transmit their signs, the correspondence conventions used to sort out system activity, the extent of the system, topology utilized as a part of the system. Networks can be categorized as following:

1. Transmission media based networks like wired networks (communication takes place through wires) and wireless networks (communication takes place wirelessly).
2. Network Size based networks like MAN, LAN and WAN.

1.2.7 Display

Use LCD s rather than LEDs, as they consume less energy. LCD screens are also easier to eyes[16].

2. Ant Colony Optimization

ACO[17] are the swarm optimization algorithms that are inspired from ant colonies behavior. It was originally proposed by Dorigo. It is based on foraging behavior of ants. Ants move in forward direction in the search of the food via shortest path. When ants get food then it starts dropping pheromes[18]. The other ants find the shortest path to get the food i.e. optimum solution. Mathematically, it can be written as below:

The transition a_{ij}^{k} (y) of the kth ant moving from node i to node j at time step y is defined as:

$$a_{ij}^{k}(y) = \{ (\zeta ij ((y)) \zeta) / (\sum [\zeta ij ((y)) \zeta \square)$$

When an ant has traversed a path within the time step y, the ant retraces its path and deposits pheromone on every edge of the path. The concentration of the deposited pheromone on the edge (i, j) is inversely proportional to the length of the path[19]. That is,

$$\frac{1}{Z^{k}(t)})$$

The total intensity can be seen as below:

$$\zeta ij ((y + 1) = (1 - \beta_{eva}) * ij ((y) * \sum \zeta ij ((y)) \zeta$$

3. Virtual Migration

The idea of Virtual Machines (VMs) as shown in Figure 1 is connected to diminishing energy utilization as it essentially decreases the rate of idle power in the general base[20].

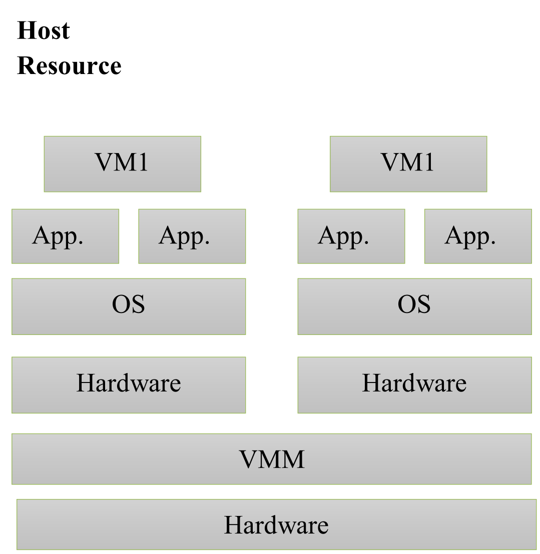

Figure 1. Virtual migration concept.

Expecting that the datacenter contains various physical servers with the asset limit RC = {RC1, RC2, RC3… ..RN} is the measure of limit for every asset (CPU, memory, plate). There are a predefined set of VM sorts VT= {VT1, VT2, VT3, VT4… ..VTN}[21]. For one period, the datacenter gets MapReduce employments from multi-client. Given the arrangement of VMs connected with every client alongside their asset prerequisites, VM situation finds the most productive pressing formulas of these VMs[22] on physical servers (PS) and is given by taking after:

$$PS \bullet \sum RC * \sum VT / \sum RC$$

4. Proposed Methodology

The methodology of proposed work is shown below using steps

1. Set of task_t and servers are taken as information. The planning of tasks_t to the servers and the server centre vitality utilization is given as yield of the calculation. The clients will ask for computing different sorts of errands relying on the RAM with time.

2. Each task_t may fall under a specific errand sort like perusing record substance, overhauling information, transferring documents, downloading programming, and so forth. In light of the kind of errand chose, the preparing time fluctuate.

3. Now, introduce server centre. A server centre is used to PC frameworks and related segments, for example, information transfers and capacity frameworks. The quantity of direction in every errand is acquired. Vitality slant is computed for every assignment of various sorts in every server with the assistance of handling time.

4. Task allocation is done in such a way that most-efficient-server gets the tasks first. If, average job density is low, i.e. then it will work on same system and if the value is high then it will move to next one. Number of active servers among the set of available servers is reduced.

The algorithm as shown in Figure 2 follows a scheduling and it will divide the jobs into servers and start execution of the job tasks.

The results are being evaluated using job computation job completed and energy consumption. Energy consumption is calculated by using the number of instructions and the energy slope.

Lemma 1: Energy and Power Efficiency Algorithm

1. START
2. Initialize = servers
3. Input= server names
4. Input= execution time
5. Do
6. Check power= p
7. Check load= l
8. Check jobs= j
9. If no. of Jobs >5
10. Den repeat step 8.
11. Generate job, j specifications
12. Again repeat step 6, 7, and 8.
13. Initialize ACO Algorithm{
14. choose initial_population
15. evaluate each individual's fitness_f
16. determine population's average_fitness
17. repeat
18. select best-ranking_individuals to reproduce
19. evaluate each individual's_fitness
20. determine population's average_fitness
21. until terminating_condition
22. the desired_fitness or enough generations have passed)
}
23. Check number of jobs executed
24. Sop

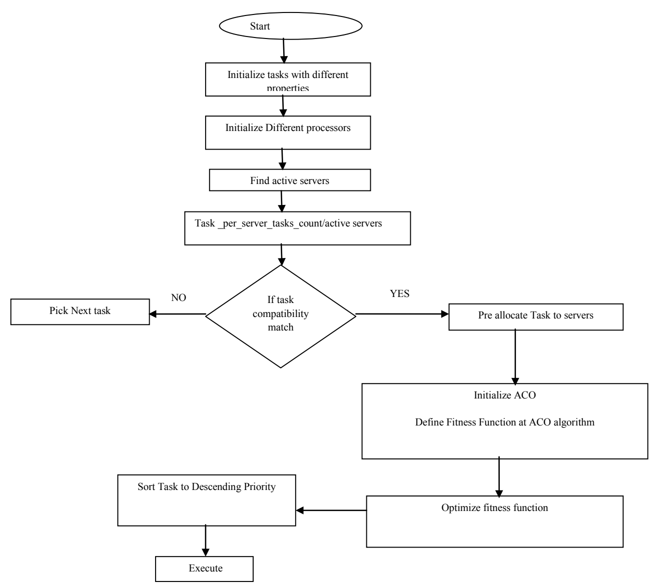

Figure 2. Proposed methodology flowchart.

5. Results and Discussion

This section describes the results obtained after the execution of the proposed work. Table 1 is for Power consumption and migration parameters that are shown for maximum value and lower threshold. For Power consumption, the maximum value is 13 and the lower threshold is 30. Similarly, for migration, the maximum value is 6 and the lower threshold is 60. Table 2 shows the comparison of power consumption, Throughput, SLA violation and Response time. These parameters are compared on the basis of Normal value, optimized value and lower threshold. The Graph for the same is shown in Figure 3.

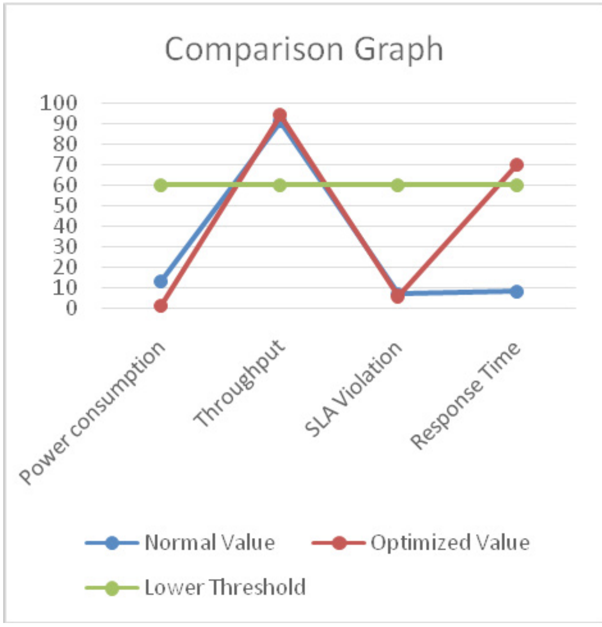

Figure 3. Comparison Graph among various network metrics.

Table 1. Parameters w.r.t Lower threshold

Parameter	Maximum Value	Lower Threshold
Power consumption	13	60
Migration	6	60

Table 2. Parameters comparison with Normal value w.r.t optimized value

Parameter	Normal Value	Optimized Value	Lower Threshold
Power consumption	13	1.5	60
Throughput	91	94	60
SLA Violation	7	6	60
Response Time	8	70	60

6. Conclusion

The Green Cloud computing takes care of the issue of a global warming change by giving eco-accommodating environment. The primary point of Green Cloud computing is to diminish the vitality devoured by physical assets in server center and spare vitality furthermore expands the execution of the framework. Additionally, each of server might be allotted to perform distinctive or comparative capacities. The tasks to be scheduled in the cloud are with 4 major attributes such as Power consumption, Throughput, SLA Violation and Response Time. So, data center resources need to be managed in resource -efficient manner to drive Green Cloud computing has been proposed in this work using Virtual machine concept with the Ant colony optimization algorithm. The results evaluation has been done in Table 1 and 2.

References

1. Ye K, Huang D, Jiang X, Chen H, Wu S. Virtual machine based energy-efficient data center architecture for cloud computing: A performance perspective. IEEE/ACM International Conference on Green Computing and Communications, China; 2010. p. 171–78.
2. Jayasinghe D, Pu C, Eilam T, Steinder S, Whalley I, Snible E. Improving performance and availability of services hosted on IaaS clouds with structural constraint-aware virtual machine placement. IEEE International Conference on Services Computing, India; 2011. p. 72–9.
3. Man CT, Kayashima M. Virtual machine machine placement algorithm for virtulized desktop infrastructure. Proceedings of IEEE CCIS, Japan; 2011. p. 334–7.
4. Lin C, Liu P, Wu J. Energy-aware virtual machine dynamic provision and scheduling for cloud computing. IEEE 4th International Conference on Cloud Computing, Taiwan; 2011. p. 736–7.
5. Jadeja Y, Modi K. Cloud computing-concepts, architecture and challenges, International Conference on Computing, Electronics and Electrical Technologies (ICCEET), India; 2012. p. 877–80.
6. Buyya R, Yeo CS, Venugopal SJ, Broberg B, Brandic I. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future of General Computer System. 2009; 25(6):599–616.
7. Yamini B, Selvi SDV. Cloud virtualization: A potential way to reduce global warming. Recent Advances in Space Technology Services and Climate Change (RSTSCC); 2010. p. 55–7.
8. Witkowski M, Brenner P, Jansen R, Go DB, Ward E. Enabling sustainable clouds via environmentally opportunistic computing. IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom), China; 2010. p. 587–92.

9. Cavdar D, Alagoz F. A survey of research on greening data centers. IEEE Global Communications Conference (GLOBECOM), Turkey; 2012. p. 3237–42.
10. Ding Y, Qin X , Liu L, Wang T . Energy efficient scheduling of virtual machines in cloud with deadline constraint. Science Direct. 2015; 50:62–74.
11. Zhu X, Laurence T, Yang Y, Wang HCJ, Yin S, Liu XC. Real-time tasks oriented energy-aware scheduling in virtualized clouds. IEEE Transactions on Cloud Computing. 2014; 2(2):168–88.
12. Tsai J-T, Fang J-C, Chou J-H. Optimized task scheduling and resource allocation on cloud computing environment using improved differential evolution algorithm. Computers & Operations Research. 2014; 40(12):3045–55.
13. Li Q, QinfenHao Q, Xiao L, Li Z. Adaptive management of virtualized resources in cloud computing using feedback control. First International Conference on Information Science and Engineering, China; 2009. p. 99–102.
14. Walsh WE, Tesauro GJO, Kephart K, Das R. Utility functions in autonomic systems. ICAC '04: Proceedings of the First International Conference on Autonomic Computing. IEEE Computer Society, USA; 2004. p. 70–7.
15. Li J, Qiu M, Niu JW, Chen YU, Ming Z. Adaptive resource allocation for preempt able jobs in cloud systems. 10th International Conference on Intelligent System Design and Application, Delhi. 2011. p. 31–6.
16. Shi JY, Taifi M, Khreishah A. Resource planning for parallel processing in the cloud. IEEE 13th International Conference on High Performance and Computing, USA; 2011. p. 828–33.
17. Dorigo M, Blum C. Ant colony optimization theory: A survey. Theoretical Computer Science. 2005; 344(2):243–78.
18. Tawfeek M, El-Sisi A, KeshkA, Torkey F. Cloud task scheduling based on ant colony optimization. The International Arab Journal of Information Technology. 2015 Mar; 12(2):129–37.
19. Mishra R, Jaiswal A. Ant colony Optimization: A Solution of Load balancing in cloud. International Journal of Web and Semantic Technology (IJWesT). 2012 Apr; 3(2):33–50.
20. Kaur P, Rani A. Virtual machine migration in cloud computing. International Journal of Grid Distribution Computing. 2015; 8(5):337–42.
21. Chawda M, Kale O. Virtual machine migration techniques in cloud environment. A Survey. International Journal for Scientific Research & Development (IJSRD). 2013; 1(8):1–4.
22. Jiuxing L, Panda DK, Wei H, Qi, G. High performance virtual machine migration. IEEE International Conference on Cluster Computing; 2007.
23. Rohini V, Natarajan AM. Comparison of genetic algorithm with particle swarm optimisation, ant colony optimisation and Tabu Search based on University Course Scheduling System. Indian Journal of Science and Technology. 2016 Jun; 9(21):1–5.
24. Silambarasan K, Ambareesh S, Koteeswaran S. Artificial bee colony with map reducing technique for solving resource problems in clouds. Indian Journal of Science and Technology. 2016 Jan; 9(3):1–6.