

Reducing Energy Consumption in Green Cloud Computing

Lestari Margatama

Faculty of Information Technology, Universitas Budi Luhur, Jakarta, Indonesia

lestari.margatama@budiluhur.ac.id

Received: 27th January 2021, Accepted: 9th March 2021, Published: 30th April 2021

Abstract

Wide-ranging cloud data centres are required for the provision of numerous software solutions across a wide range of industry verticals. Power usage is expected to rise as cloud data centres ramp up operations and expand their footprints. Here, a single-objective optimization is accomplished through the application of a Simple Genetic Algorithm (SGA). The Artificial Neural Network (ANN) uses the genetic optimization strategy to estimate the existing resources based on their features and role characteristics, and Dynamic voltage frequency scaling (DVFS) has also been utilised to reduce energy usage. The multi-objective system is empowered to reduce energy consumption because of the overall recommended method. OpenStack simulation tools were used in this context to arrive at an accurate conclusion prior to implementing on the cloud. Open Stack provides a single computer node with the ability to model and simulate a data centre. Energy usage and virtual machine migration were used to evaluate the results of the project.

Keywords: *Cloud computing, Efficiency, Energy consumption, Virtual Machine*

Introduction

As cloud computing eliminates the initial investment expenditures, the number of cloud customers has grown tremendously. It is possible to dynamically deliver hardware, software, or services via the internet using a pay-as-you-go paradigm.

Cloud service providers and their customers can both profit from a more efficient use of resources. As the processing demands of future jobs change, it is possible to dynamically lease and release new types of VMs. The need for various resources and their features is uncertain in the cloud environment. Because of the system's fundamental qualities and requirements, this is the case. The run time is the if all user activities are correctly assigned to the infrastructure. Rising energy use, on the other hand, drastically reduces available resources. Because of this, resource management and cloud planning are critical to reducing energy use and implementing the solution more quickly. Because of the ongoing monitoring of the resources, this element of the work is necessary. The preparedness of cloud users and cloud vendors for configuration is therefore crucial to minimise makeover and power usage. The cloud model is based on the idea that the system's virtual machines are computerised (VMs).

Because of the massive rise in energy costs, cloud computing has shifted its focus to reducing energy use. The reduction in energy consumption will be great if the working frequency can be effectively modified. The many heterogeneous servers can be employed in the cloud computing process and each server contains a large amount of hardware (control unit, memory unit and network devices). The energy consumption of these servers can be controlled by changing the frequency in response to operational loads. If a task requires 350 MHz, for example, a virtual machine's maximum frequency is 50 or 75 MHz (VM).

In order to save energy, many virtual machines (VMs) can do similar tasks if the frequency of operation is less than 400 MHz. The allocation of power in cloud data centres is just one of several concerns that need to be addressed. To reduce energy usage, dynamic voltage frequency scaling (DVFS) is employed to allow the VMs to run at varied frequencies and voltages. Its power usage is directly proportional to its frequency.

Review of Literature

It was predicted in [1] by Sathya Sofia et al that the competing objectives of making and consuming energy will be pursued using the semi-domination optimization technique. Power reduction, real-time scalable resource monitoring, and real-time scheduling are among the most pressing concerns of CPSs, according to Chen et. al [2]. For these concerns, they proposed a two-tier algorithm architecture that relies on cloud computing CPSs.

In order to address the problem of multi-objective work scheduling in cloud computing systems, Laith et.al[3] developed a new elite-based differential evolution hybrid antlion optimization method. Analyses of synthetic and real trace data were performed with the CloudSim toolkit. Response time (CPU), mismatch, and makespan were calculated to improve the algorithm. Zhang et al[4] discussed the issue of virtual machine (VM) energy saving solutions on a cloud computing host that is overburdened. Because of their research into the energy implications of VM migrations, they developed energy efficient VM selection methods based on greedy algorithms and dynamic programming.

An examination of current cloud simulator surveys by Ismail et al. [5] led to a cloud simulation that was constrained to only using electrical power. Finally, the selected simulators are analysed from three perspectives: their purpose and definition; energy consumption; measuring parameters and power management; and power management.

A dynamic voltage and frequency strategy, rather than a static integration threshold or a non-power-aware technique, was found to waste less energy. As a result, customers will be able to get even better service at a lower cost. Sharma et al[7] stated that green computing has shown to be an advanced technology after overcoming all of the standard cloud service issues. The GreenCloud Simulator provides two integrated power models that do not rely heavily on the energy-saving component of the system.

Green computing methods were studied by Shuja et al[8] in the context of current IT tools in our community. This focuses on the trade-off between green computing and high-performance activities, as well as effective techniques for doing so. They also looked at the problems that will inevitably arise as IT grows in terms of its ability to conduct green operations effectively.

An intelligent, green energy-efficient scheduling technique was proposed by Lei et al. in a data centre directive that aims to increase the usage of renewables and the satisfaction of tasks while reducing system running costs. Analysis of data centre resource provision and temperature control was conducted by Pedram et al. [9].

Computer system architecture, such as hardware, operating system, virtualization, and data centres, could be classified according to Beloglazov et al [10], who proposed taxonomy of energy-efficient architecture.

The traditional data centre was the focus of work by Quan et al. [11] when looking at energy management. To save energy, the allocation of resources, consolidation of workload, and frequency adjustment were implemented. The combined power use of a machine was measured by Castane et al. [12] using power metre devices. The other way is to utilise multimeters and the motherboard directly to calculate the variables needed to analyse the complete piece of work. The academics Zheng and Cai suggested a load delivery approach that is both efficient and cost-effective. This model can dramatically lower OSPs' expenses as found by Zheng et al.[13] by taking into account the uncertainty of electricity market fluctuations and the use of energy-efficient technologies in each data centre.

Using a time-conscious model for energy efficiency, Wang et al. [14] found a solution to the problem of assigning virtual machines to servers for reduced traffic and favourable traffic design conditions.

There are less active switches and coordinated flows in this method. energy consumption estimates for virtual machines are reliant on the virtual machine's processor events. Instead of cpu time, the Energy-Credit Scheduler (ECS) has been implemented to prepare a VM according to energy allocation. However, it relies solely on energy consumption, and not on the energy consumption of other components. Following this, a number of researchers presented approaches combining ANN with optimization methods in order to get optimal values while minimising resource use. A multi-target optimization approach is used to achieve these numerous, seemingly incompatible goals.

The problem of optimization is multi-objective (VEPSO). Multi-target optimization issues become a single objective optimization problem after the proper weights were added. Although many academics have attempted to minimise either the running time or the energy consumption of the cloud data centre, the focus has not been on reducing both the output and the power consumption simultaneously.

Research Methodology

System Architecture

The High-Level Cloud Architecture depicted by Figure 1. The Kilo Watt API (KWAPI) energy management system can manage OpenStack clouds, as shown in Figure 2.

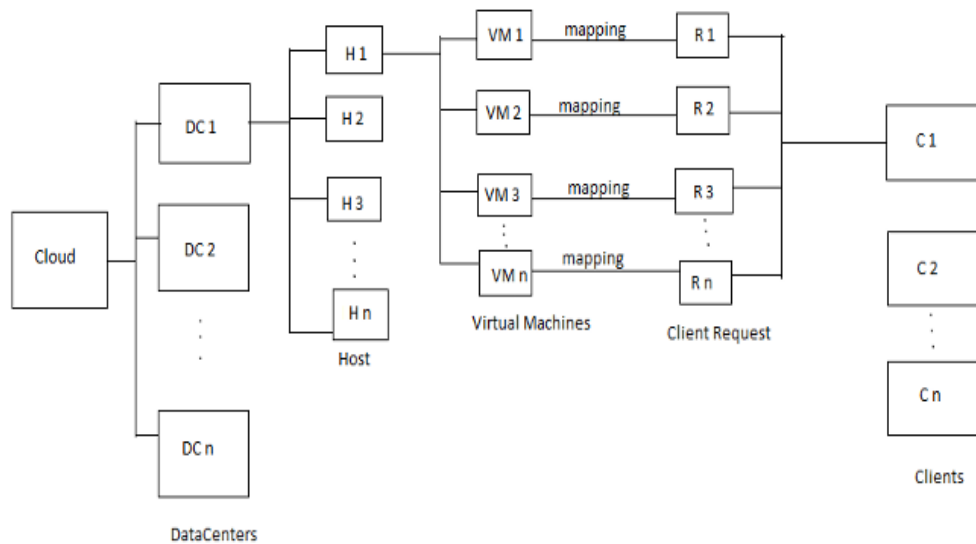


Fig 1: Cloud Architecture

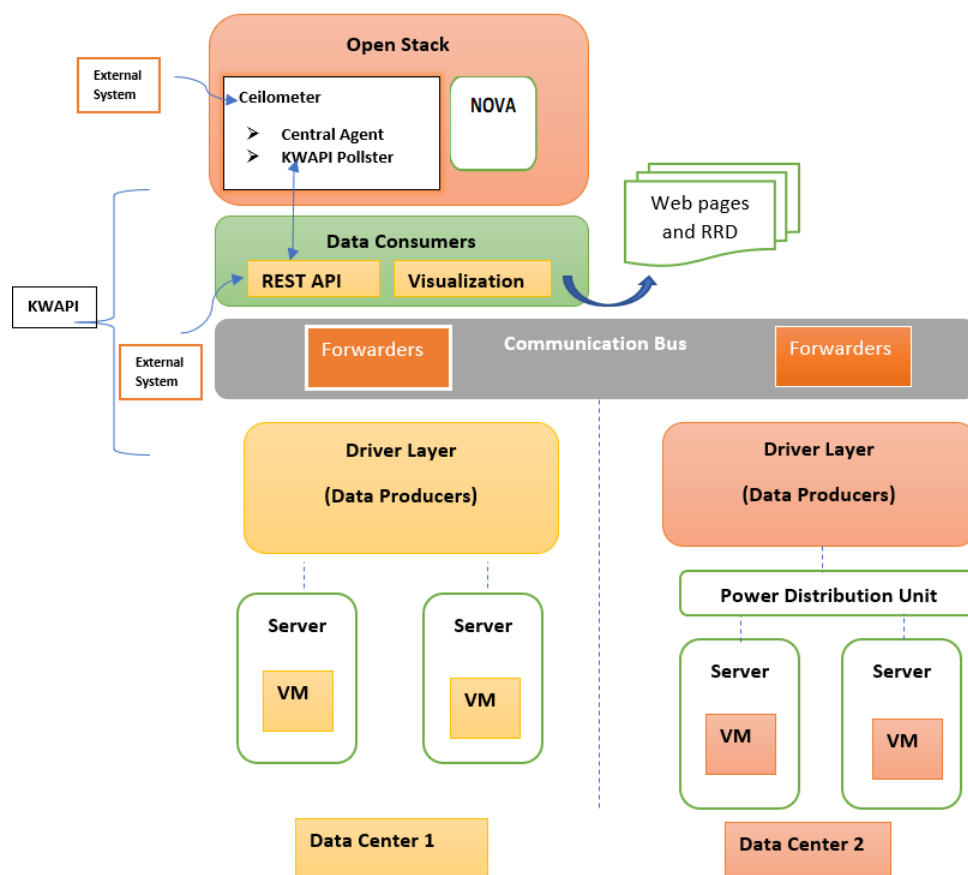


Fig 2: KWAPI Architecture

Data Consumers: A data consumer recuperates and manages driver-supplied facets by bus. The information is presented to consumers by other services such as ceilometers and visualization tools. Consumers may subscribe to all or a subset of producers by using a system of prefixes. A consumer checks the signature, extracts the material, and processes the data when he receives a message. In general, KWAPI offers the REST API (used as Ceilometer interface) and the Visualization User to two customer data users.

Driver Layers: Driver Layer Drivers are threads that are adapted to the OpenStack configuration by a driver manager with a group of file metrics loaded from the file. These parameters are used to search the meters and decide for the sensor ID to use in the collected parameters (e.g. port and IP address). The dimensions obtained by a driver are expressed as dictionaries with a limited footprint and this can be interpreted easily. Dictionary sizes differ with the amount of regulated fields set (i.e. if message signing is activated)

Interface with Ceilometer: Ceilometer 's central officer is used for transfers and storage energy measurements into the Ceilometer servers through a particular pollster (i.e. KWAPI Pollster). We ask the user to access the REST API data and to transmit (kWh) and measure counters not attached to a particular occupant because a server may concurrently host multiple clients.

Internal Communication Bus: ZeroMQ is used by KWAPI, and utilizes a high-speed C++ textbox, but no weak messaging system. A large array of bus genres, including IPC, TCP and cross-thread communication, is provided by ZeroMQ. It is easy to switch from mode to mode. There are also numerous design settings, such as subscribe / publish and request / answer. Publishers are drivers, and subscribers are data consumers. If no consumer of the data accepts the data from a given driver, no information is sent via the network.

Proposed Algorithm

The first move is to get the overloaded computer in our proposed algorithm. If there is an overloaded machine, it wakes up the standby physical machine and allocates resources for it. Another thing that can be done is to identify the low load machine. If so, pick it as a physical relocation unit and transfer it. Switch off the physical migration system upon completion of the mission. In Figure 3, the process was shown as a flowchart.

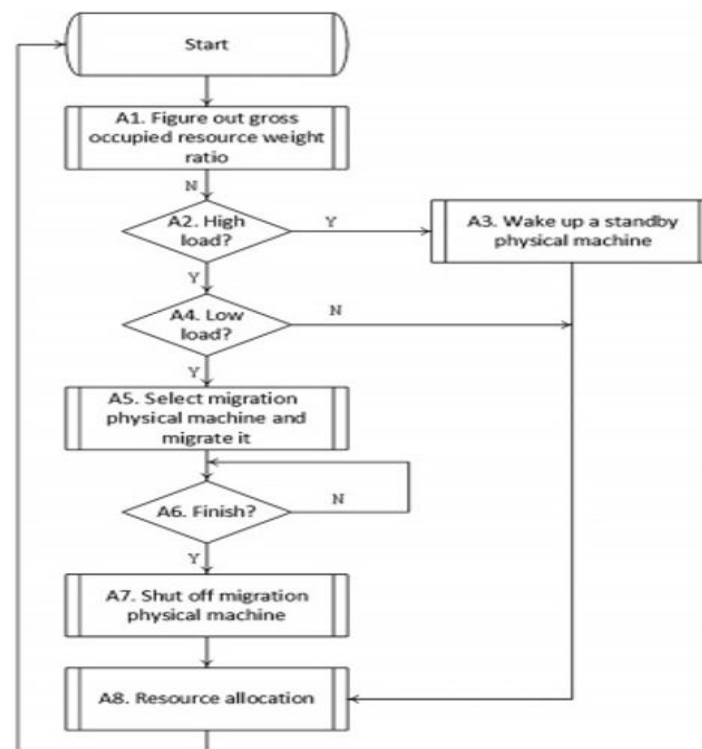


Fig 3: Host Selection

VM Migration: In this phase, occupying the resource weight is calculated. Based on it, if there is any need of VM migration, then a migration source and a migration target machine is selected. For then migration difference is calculated. Based on it, migration VM is selected and the process of migration is completed. This process has been depicted by a flowchart shown in Figure 4.

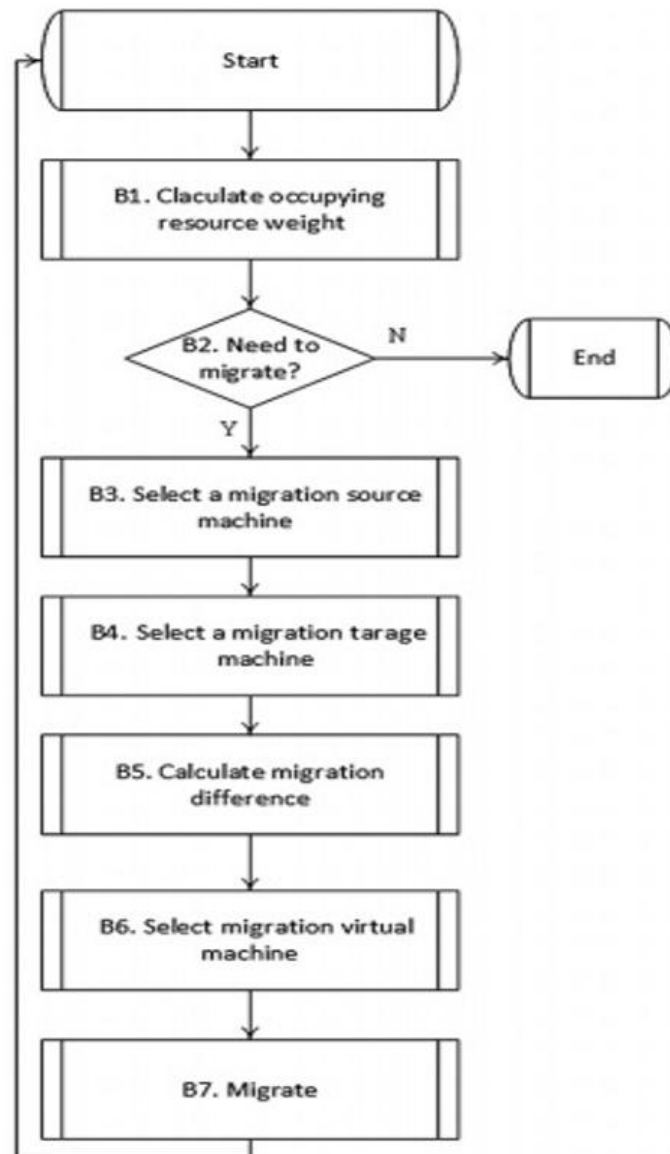


Fig 4: VM Migration

In order to ensure standard of service parameters and rising energy usage, the amount of VM should be sufficient. If there is more VM, it consumes more resources, or if there is fewer VM, the Service Level Agreement (SLA) will be broken and system performance will be reduced. Therefore, anytime over 70% of VM becomes overwhelmed, the developer installs a fresh Server. Equation 1 corresponds to this state.

$$\text{Number of overloaded VM} > 0.7 * \text{total VM} \quad (1)$$

← create a new VM

Once the work is done, the virtual machine is freed, so some virtual machines must be shut off in order to achieve optimum results. Hence

$$\text{Number of overloaded VM} < 0.3 * \text{total VM} \quad (2)$$

Therefore, a virtual machine is overloaded if it holds equations (3), (4) and (5) true:

$$TU_M > TH_M \quad (3)$$

$$TU_R > TH_R \quad (4)$$

$$TU_B > TH_B \quad (5)$$

TUM stands for the actual MIPS use, while THM stands for the MIPS threshold. Actual RAM utilisation (TUR) and threshold RAM utilisation (THR) are two different concepts. TUB is the actual Bandwidth utilisation, and THB is the threshold value of Bandwidth utilisation.. A virtual machine's memory and bandwidth will exceed the maximum limit if it is overtaxed while running on MIPS in real-world applications.

Continuous VM migration may take place in order to utilize resources effectively. To this end, it is important to estimate the resource usage by hosts. Thus, each host should be computed on the basis of Equation 6, 7 and 8 for its RAM and MIPS and bandwidth usage. Divide the overall usage of RAM by all simulated computers through the use of real devices in the System. The usage of MIPS and bandwidth is often measured in a specific fashion for any host. The average host use of RAM, MIPS and bandwidth provides an estimate of the total host use of resources (Equation 9).

$$TU_R = \frac{\sum_{j=0}^m R_{vj}}{R_{pm}} \quad (6)$$

$$TU_M = \frac{\sum_{j=0}^m M_{vj}}{M_{pm}} \quad (7)$$

$$TU_B = \frac{\sum_{j=0}^m B_{vj}}{B_{pm}} \quad (8)$$

$$TUR = [TU_R + TU_M + TU_B]/3 \quad (9)$$

Where,

TUR = Total Utilization of the Resource

TUR = Utilization of RAM (Memory)

TUM = Utilization of MIPS (CPU)

TUB = Utilization of Bandwidth (Network)

For VM migration, following condition must satisfy,

$$TUR > TH \quad (10)$$

Where, TH is the resource use threshold value. If equation 10 is true, VM is migrated to a balanced server where the balanced server usage is within the bottom and top threshold ranges.

As the cumulative days preceding the day, the amount of VMs is set. In the next step, the overloaded amount of virtual machines in the VM list will be counted. If a total overloaded VM is more than 70%, a new VM will be created. This new VM increases the overall VM count by 1. Conversely, while the overloaded VM amount falls below 30% the lower loaded Cpu falls excluded from that Cpu after the load is transferred and the amount of overloaded VM is therefore decreased by 1.

Results and Discussion

Performance Analysis: The efficiency of our proposed algorithms has been evaluated in this portion. Experiments in an incredibly complex environment, including cloud, can not be carried out repetitively and confidently. This is why, before implementing our own algorithms in actual cloud, the OpenStack simulation resources have been selected. Open Stack supports data center modelling and simulation on a single computer node.

Table 1: Energy consumption

No. of Cloudlets	Energy Consumption (KW)	
	Threshold Method	Proposed Method
1000	90	86
1500	95	91
2000	120	116
2500	140	136
3000	160	136
3500	185	181

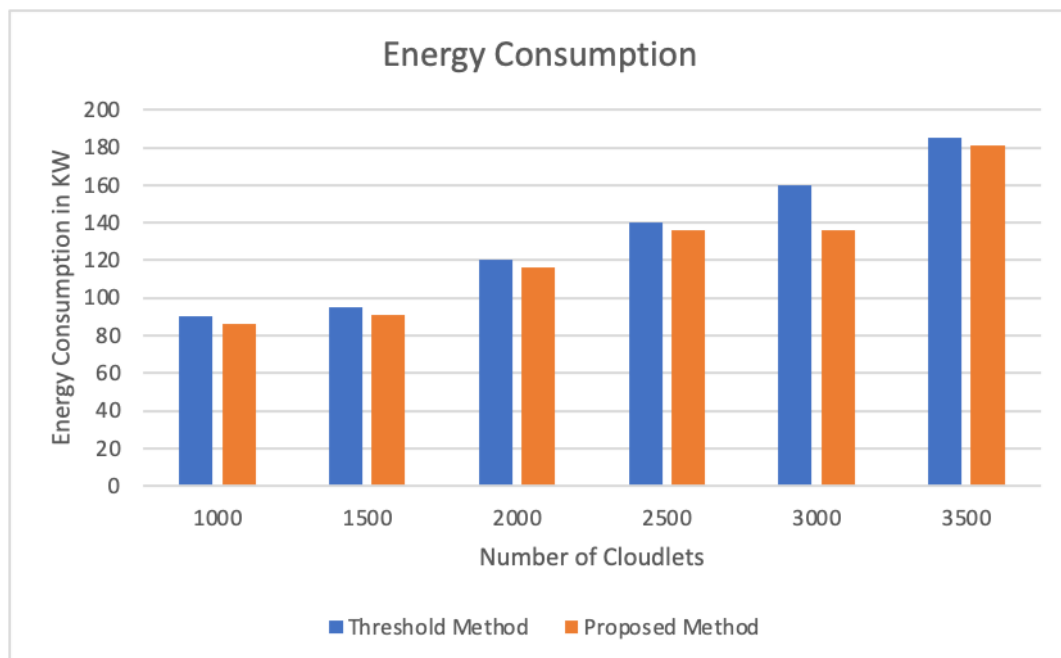
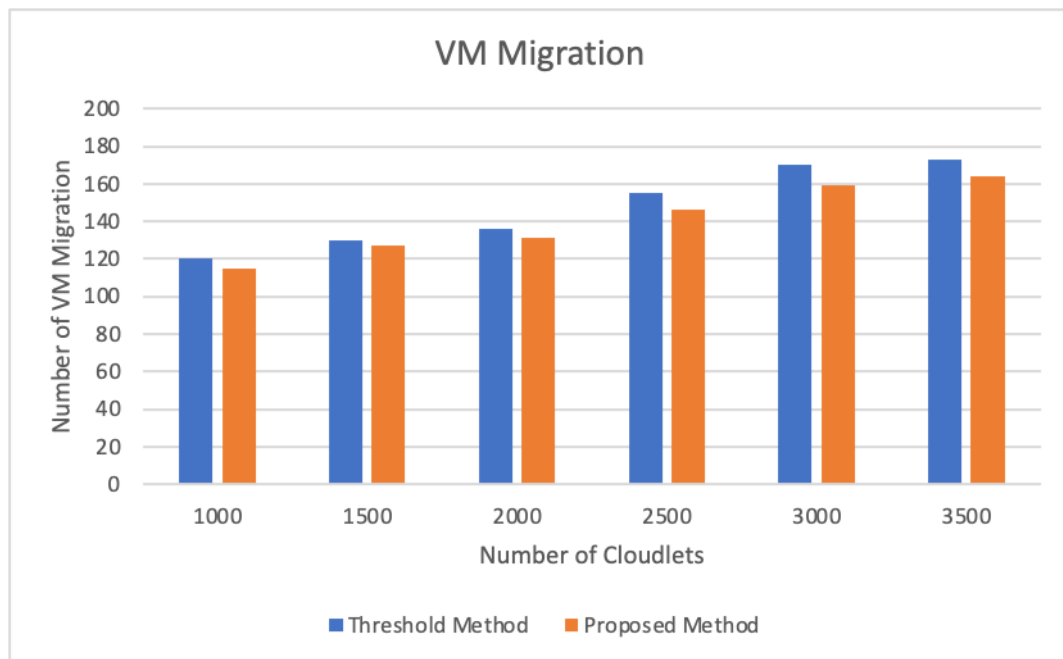


Fig 5: Energy Consumption Comparison

Table 2: VM Migration

No. of Cloudlets	Number of VM Migration	
	Threshold Method	Proposed Method
1000	120	115
1500	130	127
2000	136	131
2500	155	146
3000	170	159
3500	173	164

**Fig 6: VM Migration Comparison**

This article includes the following major contributions:

- The single-objective optimisation of energy production and results is achieved using the Simple Genetic Algorithm (SGA).
- Included in the optimisation process is the dynamic voltage frequency scaling (DVFS) to reduce energy consumption.
- A genetic optimization model uses the Artificial Neural Network (ANN) to determine the resources available on the basis of their characteristics and role characteristics.
- NSGA-II enables the multi-objective design to concurrently reduce energy usage.

For several green cloud scientists, the overall energy usage of the datacenter is the most important efficiency metric. The energy absorbed in the 1000, 1500, 2000, 2500, 3000 and 3500 cloudlets by the suggested algorithms and thresholds as shown in Table 1 and Figure 5. In each diagram, the X-axis is the number of cloudlets, and the Y-axis is the energy used in the KW / hour. The total maximum energy usage corresponds to 3–6 percent higher than the estimated energy use. In general, with all datacenter scales and charges, the proposed algorithm improved data-center energy consumption levels far beyond threshold-based method.

Another characteristic of several green cloud researchers is the total number of VM migrations inside the datacenter. A Table no 2 and figure 6 depicts the number of VM migrations in a datacenter comprised of 1000, 1500, 2000, 2500, 3000, and 3500 cloudlets per DBDV, Threshold-based method, Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO) algorithms. The X-axis of each graph reflects the datacenter load by 500 times the amount of clouds, while the Y-axis is the amount of the VM migrations. The maximum VM migration is 4-6 percent greater than the expected VM migration in the algorithm.

Conclusion

The foremost objective of this work considered in the paper is to reduce the make-up and energy usage of the cloud infrastructure. Through optimally controlling the frequency and tension of the virtual computer utilizing the DVFS method, the energy usage was efficiently regulated. The optimization process also used the artificial neural network to forecast the suitable resource on the basis of the specific terms and resources accessible. The primary step was to optimize the single objective and results indicate that by using DVFS in the optimization process, energy consumption was effectively reduced. Therefore, the issue of optimization was translated as a multi-objective, with the help of ANN and without help from ANN, and non-domination solutions were obtained with NSGA-II. As per statistical comparison it can be observed that 3-6% reduction in energy consumption and number of VM migration is reduced by 4-6 % using proposed method when compared to existing method.

In future research it is expected to define the API support for measuring power consumption and the APIs that support OpenStack to minimize power consumption.

References

- [1] Sathya Sofia, A., GaneshKumar, P. Multi-objective Task Scheduling to Minimize Energy Consumption and Makespan of Cloud Computing Using NSGA-II. *J Netw Syst Manage* **26**, 463–485 (2018).
- [2] Y. Chen, X. Chen, W. Liu, Y. Zhou, A.Y. Zomaya, R. Ranjan, S. Hu, Stochastic scheduling for variation-aware virtual machine placement in a cloud computing CPS, *Future Generation Computer Systems* (2017).
- [3] Abualigah, L., Diabat, A. A novel hybrid antlion optimization algorithm for multi-objective task scheduling problems in cloud computing environments. *Cluster Comput* (2020).
- [4] K. Zhang, T. Wu, S. Chen, L. Cai and C. Peng, "A New Energy Efficient VM Scheduling Algorithm for Cloud Computing Based on Dynamic Programming," 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud), New York, NY, 2017, pp. 249-254
- [5] Ismail, A. Energy-driven cloud simulation: existing surveys, simulation supports, impacts and challenges. *Cluster Comput* (2020).
- [6] Ahmad, Bilal, Maroof, Zaib, Mcclean, Sally, Charles, Darryl and Parr, Gerard (2019) *Economic impact of energy saving techniques in cloud server*. Cluster Computing. ISSN 1386-7857
- [7] Shivangi Sharma, Gaurav Sharma, *Impact of energy-efficient and eco-friendly green computing*, International Journal of Computer Applications (0975 – 8887), Volume 143 – No.7, June 2016
- [8] Shuja, J., Ahmad, R.W., Gani, A. *et al.* Greening emerging IT technologies: techniques and practices. *J Internet Serv Appl* **8**, 9 (2017).

- [9] E. Pakbaznia and M. Pedram, "Minimizing data center cooling and server power costs," Proc. of International Symposium on Low Power Electronics and Design, pp. 145-150, Aug. 2009.
- [10] Anton Beloglazov, Rajkumar Buyya, Young Choon Lee, and Albert Zomaya, "A Taxonomy and Survey of Energy-Efficient Data Centers and Cloud Computing Systems," Advances in Computers, Marvin V. Zelkowitz (editor), Volume 82, Pages: 47-111, Academic Press, USA, 2011.
- [11] Quan, D.M., Mezza, F., Sannenli, D., Giafreda, R.: T-Alloc: a practical energy efficient resource allocation algorithm for traditional data centers. Future Gener. Comput. Syst. 28(5), 791–800 (2012)
- [12] Castane, G.G., Nunez, A., Llopis, P., Carretero, J.: E-mc 2: a formal framework for energy modelling in cloud computing. Simul. Model. Pract. Theory 39, 56–75 (2013)
- [13] Zheng, X., Cai, Y.: Energy-aware load dispatching in geographically located internet data centers. Sustain. Comput. Inform. Syst. 1(4), 275–285 (2013)
- [14] Wang, L., Zhang, F., Arjona Aroca, J., Vasilakos, A.V., Zheng, K., Hou, C., Li, D., Liu, Z.: GreenDCN: a general framework for achieving energy efficiency in data center networks. IEEE J.Sel. Areas Commun. 32(1), 4–15 (2014)