

Power Optimisation Model for Leveraging Cloud System

Ogechukwu M. Okonor, Dr Mo Adda, Dr Alex Gegov.

School of Computing
University of Portsmouth, Portsmouth
United Kingdom

Abstract— *Cloud computing is an emerging technology that has significantly increased over the years due to an increase in the number of internet users and its numerous benefits. It is a technology that leverages cost and considers locational challenges. The emergence of cloud computing has brought innovation and hope to many under-developed countries with a less skilled workforce, limited resources and unstable electricity supply. Although this technology is transforming information technology, it still faces some significant challenges like high energy consumption rate, confusion in adoption method and security issues. The focus of this paper is on reducing the problems with ambiguity on data centre structure and its power usage by simplifying how the application works and then provide a holistic approach to minimizing its power usage rate. This paper minimizes the energy utilization of the cloud data centre using an intelligent method to shut down, schedule and monitors the entire cloud system.*

Additionally, observing the complexity of the cloud system design, a holistic approach to calculating and measuring the precise amount of power usage during each operational cycle is vital to achieving a more efficient data centre energy model. This work proposes an improved energy model that calculates, in detail, the energy usage on the system by focusing on each system component and its contribution to the high energy consumption level of the data centre. Furthermore, the proposed model is then used to develop an algorithm that minimises the power usage in the cloud environment to an optimal low level while maintaining a high level of service level agreement.

The result obtained from this experiment shows that the developed model removed a high level of obscurity in the issue of the cause of high energy consumption in the data centre and thereby encourage cloud providers to act accordingly in reducing the effect.

Keywords – Data centre, Energy consumption, Cloud computing, Virtualization, Link Utilization and Virtual Machine.

I. INTRODUCTION

Cloud computing has become a household name with a significant link to all emerging technology. With the tripling rate of internet users and high adoption of wireless devices, it is evident that cloud technology is the future and it leverages the deadlock in the world of Information technology when it comes to acceptability and affordability. It was initiated in 2006 by Amazon [1]. Also, observing the current trend in Information Technology (IT) innovations and finding a balance technological approach to relieve companies of infrastructural stress “Cloud is a must embraced technology”.

Although it comes with a fantastic remuneration, it still has some challenges that need urgent attention. That is, the cloud system cannot

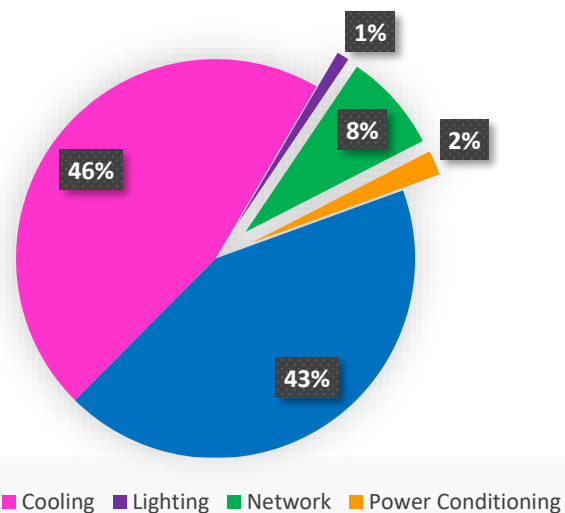
function effectively without the aid of the critical infrastructure of the data centre.

Data centre (DC) is a network of computing and storage devices that consist of core, switches, pods, servers and virtual machines (VMS) which enables the delivery of shared software application and data.

This is the ugly part of the cloud technology because the data centre is a power-hungry device which must be powered 24/7 to be able to satisfy the service level agreement (SLA) promised to clients [2].

FIGURE 1 SHOWS THE PROPORTION OF WHICH THE ENERGY CONSUMPTION LEVEL IN DC COMPONENT

Datacenter Energy Breakdown



This will also be synchronized with the energy consumption rate of the global data centre and US, which is hosting the higher part of DC activities in the Table 1. Note in the table below, the use of the following abbreviations: Estimated Annual Energy Cost (EAEC), Estimated Energy Consumed (EEC) and Energy Used by Data Centres (EUDC)

Year	Investigator(s)	EAEC (USD Billion)	EEC (TWT)	EUDC (%)	Region
2005	Miyuru et al		0.061	1	US
2006	R. Brown et al	4.5	0.089	1.5	Globally
2007	Van Heddeghem		216	1.8	Globally
2009	G.Meijer et al	30		2	Globally
2011	Google vendor	163	2.68	10.2	US
2012	Garter	106.4		12.7	
2013	Ni		91	13.2	US
2016	X. Fan et al	240			
2017	Marriott Marquis		102	14.1	US
2017	Atlanta		416	3	Globally
2019	Energy watch Manager	10.77	424	4.7	Globally

Table 1: Tabular analysis of energy consumed by DC from 2006-2019.

Furthermore, these power usage increases can be evidenced to Andrea [3], a Swedish researcher who predicted in 2016 that by 2025, the data centre would amount to the most significant ICT share of global electricity production at 33% followed by smartphones (15%). This was recently (2019) confirmed by an American researcher in an article that the data centre will be using up to 20% of available electricity in the world by 2025.

Therefore, the rest of this paper is organized as follows: Section 2 reviews the related works; section 3 proposed mathematical model for DC system; section 4 the proposed flowchart and result analysis section 5 conclusion and future work

2 Related Work

Many published works have contributed to the improvement of energy efficiency and management. Despite their contribution, according to the evidence shown in Table 1 above, there is still an urgent call for a more focused optimized solution to the data centre energy consumption rate. The continuous effect can be attributed to the increasing demand in traffic activities due to the introduction of high bandwidth requirement and Quality of Service (QoS) measures in new applications. In this context, the power consumed during runtime is always steady once it is powered on-; therefore, it does not scale with the active traffic load [4]. In 2013, Orgerie et al. acknowledge that the data centre energy consumption experiences more energy wastage than the amount of energy used by leaving computing and networking equipment powered on 24/7. The evidence suggests that only 30% of the cloud resource is used at a time, leaving 70% as a wasted resource. There are many studies on dynamic voltage frequency scaling (DVFS) technique for managing power usage, which this research work assumes is the gateway to finding a perfect solution under many approaches. Therefore, building an alternative hypothesis that uses a lightweight intelligent as it approach will achieve a more optimal result. For example, the work of (Sharma et al.) [5] focused on the single server approach, which is not realistic, with the cloud environment due to its complexity and multi-server connectivity.

Beloglazov et al. (2011) [6] did many research works on energy efficiency in network and data centre. Some of his research focuses on how to manage the energy consumption of the Computer memory (processor) using the DVFS technique.

Whereas, Sinha et al. in 2011[7], proposed a CPU utilization dynamic threshold base used for predicting and managing workload in the data

centre. However, this has some limitations as well due to some congestion and bottleneck, it experienced during the process therefore not optimally minimizing energy.

Furthermore, Isam A. et al in 2013 took a different approach entire which was considering the scheduling process involve in cloud data centre, it looked at resources where shared and scheduled in the cloud system to minimize it redundancy and overload on some of the servers indirectly saving energy. Nonetheless, this work also has some limitations because it just considered the computing resource not the network resource.

The diurnal traffic of a data centre is unpredictable and fluctuates in between the peak and off-peak periods-; because it is completely an online process with no measured procedure. This has led to the massive deployment and over-provisioning of CC resources to cope with the traffics on 24 hours basis. It has been observed by [LIU et al [8] that the average load in data centres is about 30% of the resources and this underscores the high level of resource over-provisioning. Therefore, reducing the active resources in an intelligent systematic way that the cloud DC efficiency is not sacrificed is main aim of this paper. Of course, this can save up to 70% of the energy used. It has also been shown that the line-card energy consumption in IP routers dominates the total energy used by the network to the tune of 43% of the total power of the router [Zhang et al, 9]. Additionally, the authors proposed the combination of both traffic and work load consolidation schemes. Of course, this work is not the first proposal on energy efficiency on CC but most research do not optimize both the efficiency of CC and its energy savings.

Looking at a typical data centre diagram in Figure 2 it will be easy to understand the reason behind the complexity of its connection. The diagram below represents a 3 tier Dc topology with its core, aggregation, access and server layers. The ability to synchronize the traffic movement in this DC structure in both active and non-active period is the research target.

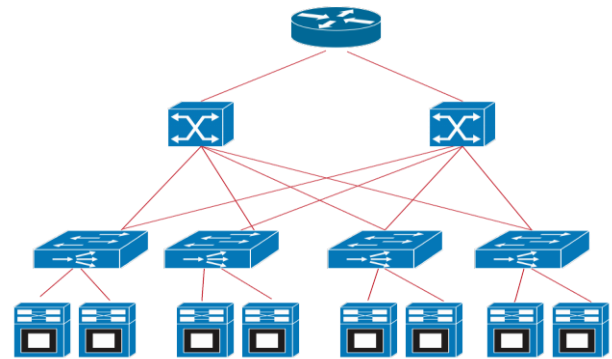


Figure 2: Data centre structure

3 Proposed Mathematical Formulation of Energy Models in DC.

This research work suggests from an evidential point of view that any technology that cannot be thoroughly measured cannot be thoroughly managed. Table 2 has the full description of the variable used in formulating this energy model.

Table 2: full description of the variables used in the formulation.

Variable	Definition
P_i	Ports
P	Power
E	Energy
μ	Utilization
S_{rv}	Server
R	Rate
N_c	Network configuration
V	Set of network device
\mathcal{E}	Set of links
S_{wi}	Switches
P_{cc}	Computing cost
V_m	Virtual machine
H	Host

This paper calculated the total Power consume in the DC which is the summation of the total computing power (both the computing and networking device) which sometimes was overlooked in other articles but will help give a better view of how minimizing its data centre activities can contribute to energy-efficient data centre. Therefore, the total computing cost is defined as:

$$P_{DC} = P_H + P_{Net} \quad (1)$$

$$P_{Hi} = P_{Svi} + P_{vw} + P_{SRi} \quad (2)$$

$$P_{vm} = \sum_{j=1}^k P_{Vmj} \quad (3)$$

$$P_H = \sum_{i=1}^n P_{Hi} \quad (4)$$

$$P_H = \sum_{i=1}^n (P_{Svi} + \sum_{j=1}^{ki} P_{Vmj} + P_{SRi}) \quad (5)$$

$$P_H = \sum_{i=1}^n (P_{Svi}(1 - \delta_i) + \sum_{j=1}^{ki} P_{Vmj} + P_{SRi}(1 - \beta_i)) \quad (6)$$

Now, knowing that switches form the basis of the interconnection fabric that delivers task requests to the hosts for processing. Energy consumption of a switch depends on a linear model where each individual port has it different line rate. The following is its content: number of ports, port transmission rate, types of switch and the deployed cabling solution based on the simple 3tie or fat-tree topology in fig 2 which this formulation is based on. The energy consumption model for switches then:

$$P_{swi} = P_{chassis} + n_{linecard} \cdot P_{linecard} + \sum_{i=0}^c n_{ports} \cdot P_r \quad (7)$$

* UtilizationFactors

Where

$P_{chassis}$: The power consumed by the switch base on hardware

$P_{linecard}$: The power consumed by the active line cards

P_r : The power consumed by the active port running at 'r' rate

Computing Server also contributes heavily to the high energy consumption rate in the data centre. It will be ideal to say that the energy consumption of a host server is proportional to the CPU utilization. Previous research has shown that an idle server consumes two-thirds of its peak-load use to keep memory, dish and input-output resources running. Its energy model is given as [10]:

$$E(A) = E_{cpu}(A) + E_{memory}(A) + \frac{E_I(A)}{0}$$

Building on this, the energy model for the cost of power used by the server will then be

$$P = P_{srv} - \sum_{i=1}^n \delta^i P_i \quad (8)$$

$$\text{where } \delta_i = \begin{cases} 0 & \text{if server } j \text{ is enable} \\ 1 & \text{if server } j \text{ is disable} \\ fi & \text{if server } j \text{ is in standby} \end{cases}$$

which can also be represented as $0 \leq \delta_j \leq 1$

Then

$$P = \sum_{i=1}^n P^i - \sum_{i=1}^n \delta^i P_i = \sum_{i=1}^n (1 - \delta_i) P_i \quad (9)$$

Therefore, to calculate the total power cost of the cloud computing device will refer to equation 1

Which will then be summing the Power on the host (equation 6) with that of the network (equation7).

$$P_{DC} = \sum_{i=1}^n (P_{Svi}(1 - \delta_i) + \sum_{j=1}^{ki} (P_{Vmj} + P_{SRi}(1 - \beta_i))) + \sum_{k=1}^m (P_{chassis} + n_{linecard} \cdot P_{linecard} + \sum_{i=0}^c n_{ports} \cdot P_r) \quad * UtilizationFactors$$

4 PROPOSED INTELLIGENT CLOUD DC FLOWCHART

Cloud data centres are composed of many computing/storages nodes (e.g. servers, CPU, memory), and network devices (network links). To save energy, the energy consumption of these components should be taken into account. Considering that an intelligent agent can execute asynchronously and autonomously. This research work deployed intelligent system into the architectural component of the data centre system and then, minimize the energy consumed in the network links, Virtual machine (VMs) allocation monitoring, CPU utilization and storage nodes.

The proposed flowchart in Figure 3 aims at reducing the number of links in the network during energy-saving periods. It has been designed in such a way that the energy-savings does not in any way affect the quality of service of the cloud customers. First, it takes an input, the full network topology and the given traffic matrix (as generated by the customers' requests), and the threshold which is determined by the operator. The threshold is based on the maximum link utilisation (MLU) of the cloud data centre resource and less than 100% of the customers' request to reduce system demand and increase response time. The idea is to make sure that energy-savings algorithm is triggered before congestion starts. The utilisation of each network link is computed, examined by the algorithm and sorted in ascending order according to respective utilisations - based on this. The removal process starts with the mobile agent. Before any link is removed, the cloud connectivity is checked. Therefore, confirming that no computing node is disconnected from the cloud infrastructure. If the maximum link utilisation of the entire cloud data centre does not exceed the pre-defined threshold, the link can be successfully removed and scheduled for energy savings. Otherwise, the next link is checked.

This continues until all the links are checked. This process is also applied to the switches, host and virtual machines in the system.

Traffic	Links	MLU	ALU
TM1	32 out of 74	78.16	5.08
TM2	33 out of 74	77.72	5.57
TM3	32 out of 74	80.67	5.62
TM4	34 out of 74	81.28	5.71
TM5	31 out of 74	82.31	5.77
TM6	31 out of 74	85.21	5.86
TM7	32 out of 74	84.27	5.78
TM8	31 out of 74	81.23	5.88

Table 3: Tabular analysis of a snapshot of a TM

Regarding the research findings above, there is a high rate of under-utilisation of the cloud platform. The introduction of an intelligent agent helps leverage the complexity of cloud environment by having full migration access to the cloud DC components without disrupting the quality of service or bridging the service level agreement. That is why this proposed mathematical model with an intelligent agent attached to it cannot be ignored. It balances the trade-off between efficiency in performance vs power usage in the cloud environment. Importantly this proposed work will be the first of its kind to monitor and manage the power usage during and after scheduling process involving the link, node and switches. This approach considers all the cloud components (i.e. the network, links, switches and traffic method) in saving energy in the cloud environment. It will efficiently manage, monitor and aid proper scheduling process with less response time and human error. Therefore, this work propels more energy-efficient cloud data centre and cloud environment at large.

Minimizing the links is just one of the stages of this research work, continuing further, the intelligent agent code will be used to regulate the virtual machine migration and monitoring under different threshold scenario. Finally, the intelligent code will be embedded into the switches to regulate the activities of the host without violation of SLA. Figure 4 and 5 is the graphical representation of the result.

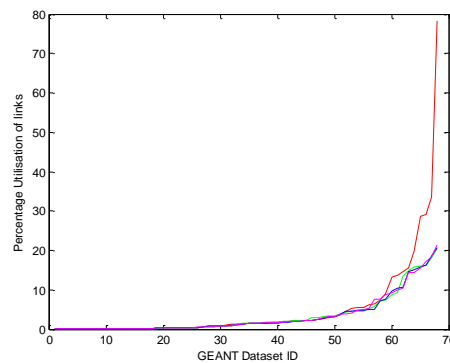


Figure 4: Percentage utilisation of each link using customers' TM

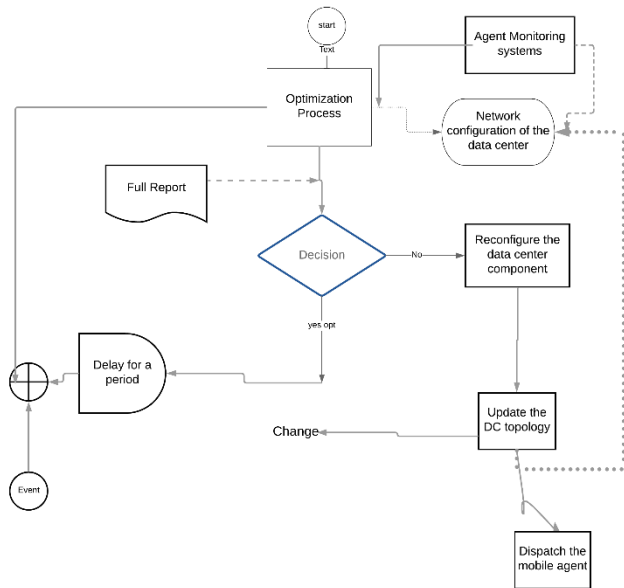


Figure. 3: Proposed intelligent agent Cloud DC flowchart

5 RESULT EVALUATION AND ANALYSIS

Before delving into the result evaluation and analysis, we analytically prove that the hybrid approach using both batch and continuous process achieved an excellent performance level regardless of its varying simulation runtime. We used CLOUDSIM platform [11] to evaluate our proposed methodology. CLOUDSIM is a new open-source platform for experimenting with cloud-related topics with an extensible simulation framework. It allows modelling, simulating and analysing cloud system design and its applications. It is easy to add more entities to the available tool on the structure to suit one's research experimental needs. It uses Java API.

In our simulation, we simulated a single data centre with 100 hosts, 70 links, 1 core CPU, MIPs varies from 1000-3000, RAM 8GB, storage 1TB and 300 VMs.

Table 3 shows the proposed method was able to remove a significant number of non-active links within 15 minutes in an hour simulation runtime.

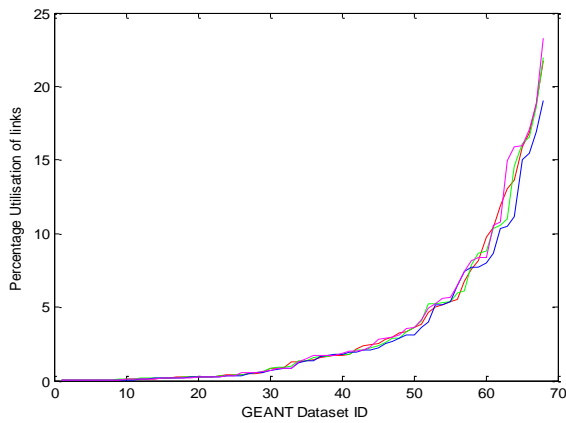


Figure 5: Percentage utilisation of each link using customers' TM for the next hour.

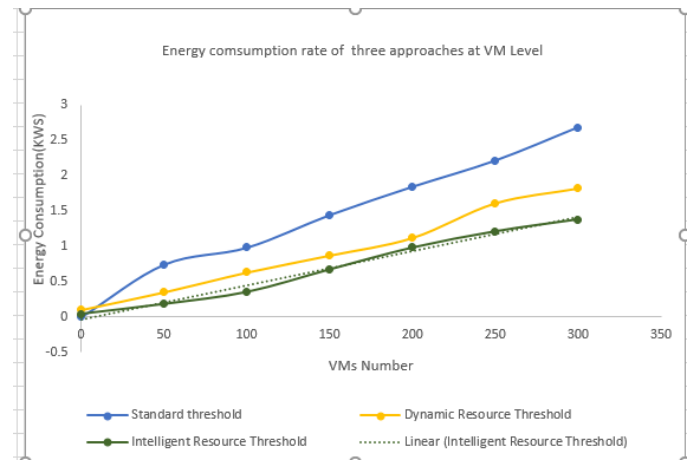


Figure 6: Comparison analysis of three different VMs threshold.

To maximise the energy savings, the intelligent agent monitors the activity of the network; then it switches some of the inactive links into sleeping mode such that energy can be saved during those periods. Figure 5 shows the amount of energy saved during the 1-hour simulation. As can be seen, a substantial amount of energy is saved by the mobile agent within the an-hour of traffic evaluation. The result figure 6 and the method of Figure 3 checked the migration process in the VMs using intelligent code to manipulate the already existing dynamic threshold. The intelligent threshold approach by this research work produced an optimal result with significant savings without violating the SLA of the system operation compare to the other standard and dynamic threshold.

6 CONCLUSION AND FUTURE WORK

Resource utilization in the cloud data centre continues to be a significant challenge, especially in the area of energy consumption, with these continue increase in the number of IoT devices and service level expectation. This research work focused on developing an online approach to mitigating the high rate of energy consumption in the cloud DC. Unlike conventional norms, energy can be saved in CC using

offline approach i.e. a strategy that focused on the night time or when traffic is less. However, such an approach does not adapt to the dynamics of the network in an online environment where the cloud system functions. In this paper, an intelligent system was used to monitor the activities of the data centre, focusing on the links during runtime. This method minimizes a good number of inactive links during low traffic time in the network. Some DC components operating on less than the set threshold was shut down for energy savings or put in sleeping mode. This is a novel approach to energy savings since it does not need to store or know the statistics of the network. Therefore, it is an online approach to energy savings. Currently, only the offline approach has been utilised, and due to the introduction of an intelligent agent system, a more significant energy savings has been recorded.

In future work, the intelligent agent code will be used to model the mathematical formulation on this paper on virtual machine migration and monitoring further, then on the switches to regulate the activity of the host without violation of SLA and Policies. This will enhance optimal energy usage in the cloud system based on the power usage efficiency (PUE) metric.

REFERENCES

- [1] Yang, Haibo and Tate, Mary, "A descriptive Literature Review and Classification of Cloud Computing Research," Communications of the Association for Information Systems: Vol. 31, Article 2.
- [2] Oro E, Depoorter V, Garcia A, Salom J. Energy efficiency and renewable and energy integration in data centres. Strategies and modelling Review. Renewable and Sustainable Energy Review 2015;429-445
- [3] Andreas Berl1, Erol Gelenbe, Marco di Girolamo, Giovanni Giuliani, Hermann de Meer1, Minh Quan Dang and Kostas Pentikousis, " Energy-Efficient Cloud Computing" Published by Oxford University Press on behalf of The British Computer Society. Advance Access publication on August 19, 2009
- [4] Ogechukwu Okonor, Mo Adda" Intelligent Approach to Minimizing Power Consumption in Cloud-Based System Collecting Sensor data and Monitoring the Status of Powered Wheelchair"Intelligent system conference 1, 2019
- [5] Sharma .A, Y Yao, L Huang- 2012 Proceeding...,2012- IEEEExplore.ieee.org
- [6] Anton Beloglazov,"optimal Online Deterministic Algorithm and Adaptive Heuristics for Energy and

- [7] R. Sinha, N Purohit, “Energy efficient dynamic integration of threshold for migration at cloud data centres”, Special issue of international journal of computer Application on Communication and Networks, no 11, pp.44-49,2011.
- [8] Liu, J., Zhao, F., Liu, X., He, W.: Challenges Towards Elastic Power Management in Internet Data Centers. Proc 2nd international workshop on cyber-physical systems (WCPS), in conjunction with ICDCS Montreal, Canada, (2009)
- [9] Berl, A., Gelenbe, E., di Girolamo, M., Giuliani, G., de Meer, H., Quan Dang, M., Pentikousis, K.: Energy-Efficient Cloud Computing. Oxford University Press on behalf of The British Computer Society (2009).
- [10] Miyuru Dayarathna, Yonggang Wen,” Data Center Energy Consumption Modeling Survey” IEEE communication Vol 18,2016.
- [11] Rodrigo N Calheiros, Rajiv Ranjan, Anton Beloglazzov” CloudSim: a toolkit for modeling and simulation of Cloud computing environment and evaluation of resource provisioning algorithms”, Wiley 2010.