A Review of Performance and Energy Aware Improvement Methods for Future Green Cloud Computing

Ali Abdullah Al-Mahruqi
Electrical & Computer Engineering Department
Caledonian College of Engineering, Sultanate of Oman

Brian G. Stewart, Brian Hainey
School of Engineering and Built Environment
Glasgow Caledonian University Glasgow G40BA, United Kingdom

Vallavaraj Athinaryanan
Electrical & Computer Engineering Department
Caledonian College of Engineering, Sultanate of Oman

ABSTRACT
With the advent of increased use of computers and computing power, state of the art of cloud computing has become imperative in the present-day global scenario. It has managed to remove the constraints in many organizations in terms of physical internetworking devices and human resources, leaving room for better growth of many organizations. With all these benefits, cloud computing is still facing a number of impediments in terms of energy consumption within data centers and performance degradation to end users. This has led many industries and researchers to find feasible solutions to the current problems. In the context of realizing the problems faced by cloud data centers and end users, this paper presents a summary of the work done, experimentation setup and the need for a greener cloud computing technique/algorithm which satisfies minimum energy consumption, minimum carbon emission and maximum quality of service.

Keywords
Cloud computing, virtual Machines, Virtualization, IaaS, hypervisors, energy consumption, performance, and energy efficiency

1. INTRODUCTION
Cloud computing has become one of the most important technologies of the present time; many organizations are enjoying cloud facilities by reducing the cost of operation and increasing the level of production and management[1],Though many countries are still hesitating in joining cloud computing due to security reasons [2], it has become almost impossible to stop the advancement of cloud technologies. An increasing number of cloud providers[3] are emerging with improved and more sophisticated services that attract many organizations to join the cloud. Most cloud computing service providers are using hypervisors to manage virtual machines (VMs) as the main methodology for server consolidation. This has proved to be one of the most important techniques to help reduce the number of physical machines in any organization. The resources of cloud computing are managed by service providers (through third parties) who offer Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS) [4]-[6]. Cloud computing can be deployed on four different models which allows flexibility of choice to the users. The four deployment models are as follow: (See Figure. 1)

- Private Cloud: This model can be deployed exclusively by an individual organization or leased by service providers dedicated to a particular organization. This model is expensive to deploy as infrastructures and services are utilized by a single organization.
- Public Cloud: This is a type of deployment model which is shared by multiple organizations. It always consists of mega-scale infrastructures with capabilities of providing services to multiple companies. It is cheap in term of cost as infrastructures and services are shared by multiple companies.
- Hybrid Cloud: This deployment model is composed of more than one deployment (private, community and public). For example an organization can choose to lease a public deployment from a service provider while maintaining in-house infrastructures (private) or leases from a service provider.
- Community Cloud: This allows sharing of infrastructures for specific organizations or communities. For example the ministry of health can deploy a community cloud for all ministries of health and hospitals around a country. It can also be managed solely by an organization or leased by service providers.

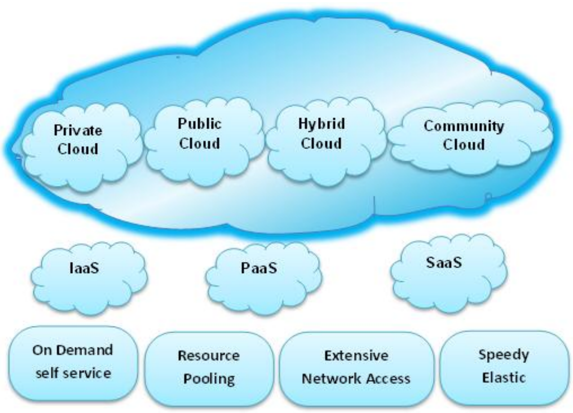

Fig.1. Cloud Deployment & Service Model

2. PROBLEM DEFINITION
Since the domain of research in the area of cloud computing is very wide, the main focus of this paper is on energy consumption, carbon emission and performance in aggressive virtual machines consolidation (VMC). In order to ensure quality of service (QoS) and energy efficiency, server VMs consolidation problems become the main focal point. Research carried out suggests that the lack of proper

planning in the use of cloud computing can cause serious problems [7]- [10]. While the users of cloud computing benefit from 90% energy savings and zero utilization of physical resources, delivery of services from providers to consumers still poses a challenge in terms of energy consumption and carbon emission.

Global warming has become one of the most serious problems facing today's world of technology. Data centers are also considered to be one of the major contributors to global warming whereby idle servers consume about 383.75W and 454.39W when they are in full operation mode. It also costs about $301.54 per year for emitting 7731.8 1bs of CO2. [11]. According to recent statistics, energy consumption is dramatically increasing every year. In the Asian specific region, it is increasing at a rate of 23% compared to a level of 16% for the rest of the world every year [12].

The initiatives of power management techniques of dynamic voltage frequency scaling (DVFS) [13] and technologies for making servers more energy efficient have been proposed by Greenpeace, the United States Environmental Protection Agency and the Climate Server Computing [14]. However the challenges are still at large as all these technologies depend on hardware which may not be controlled according to the needs of the cloud providers.

The main objective of the research in this paper is to overcome the problem faced by data centers in terms of energy consumption, performance and quality of service (QoS). Based on the existing state of the art that has been provided by different researchers, gaps have been identified to help provide better plans for server virtual machines (VMs) consolidation that will contribute to make better utilization of resources which will in turn minimize energy consumption and improve performance. A practical experiment is performed to measure CPU, Memory and Power utilization. An experiment serves and judges when a server will become under loaded, overloaded and how idle servers are compared to fully utilized servers. Based on the results, an algorithm is proposed which will later be implemented to satisfy the future trend of greener cloud computing.

3. RELATED WORK

Existing research has proposed better heuristics, techniques and algorithms that control the utilization of VMs in servers so as to make cloud computing greener, and more efficient.

Authors in[15] have proposed research on reducing energy consumption in cloud infrastructure which allows the number of physical servers (PMs) to be minimized. It allows the VMs to be moved to a limited number of servers in order to allow the concentrated physical servers to utilize maximum resources, setting other servers to a low power mode. An algorithm was developed which migrates the VMs from an empty server in order to maximize the power consumption. The principle is based on a gossiping protocol that sprints from time to time in order to spot the current situation of the VM allotment, and uses the live migration features provided by the virtual machines monitors' management to complete the migrations. The algorithm mostly considers the number of VMs operating on each server. It checks the capability of a certain server through gossiping to determine if it can allow migration to happen from a source server, which will have to make sure the destination server is not full before the migration. The method uses peer-to-peer (P2P) service to gossip in order to determine the current VMs hosted by a particular server. Though the algorithm is able to detect the number of VMs in each physical machine, it is rather difficult to judge when and how the physical machines become overloaded or under loaded, without setting up parameters such as CPU, memory and network bandwidth usage. Authors in[16]-[18] have identified various issues with performance problems in data centers based on unpredicted workload sent by cloud consumers; most of the performance issues discussed were caused by unpredicted usage of CPU, memory and network bandwidth. As that is the case, this work will focus on setting up CPU and memory parameters based on practical experiment, which might provide an exact estimation on when the CPU and memory becomes overloaded, under loaded or idle. The study in [19]-[20] suggests a similar technique of migration in the VMs from numbers of physical machines available in the data centers (DCs) and squeezing them to other physical machines which do not seem to be fully packed. Although the concepts look similar, their techniques vary from that of [15].

VM migration has been considered to be the strategy of saving energy and improving the QoS in DCs [21]-[22]. Authors in [23] have proposed VMs placement and used a modified best-fit decreasing algorithm which allow VMs to be sorted based on usage of CPU in decreasing order; this is quite different from the research in [24] which was based on millions of instructions per second (MIPS). They introduced the use of a threshold system, which is believed to be the better solution for preventing the resources from being over utilized. Higher and lower utilization have been considered rather than just higher utilization, which will always allocate the VMs to physical machines (PMs). They also suggested using dynamic migration of VMs, which will fit the changing behavior of VMs. In addition, three methods have been suggested so as to make a decision on which VMs will need to be migrated:

- Minimization of migration: a few numbers of VMs migrate so as to reduce migration overhead
- Highest potential growth: the VMs that seem to have the lowest CPU usage are migrated
- Random choice: choosing the essential number of VMs in a random manner

As per the algorithm, the authors have suggested a gossiping protocol with two threads (Active and Passive) which is similar to that of [25] which decides the number of VMs a particular PM can handle.

They [23] have also reviewed the general power model that helps to identify which components of cloud computing consume energy when the workload increases. Different components such as physical machines (PMs) with VMs, storage, and network components have been reviewed in terms of energy consumption. Though they have not considered all of the above factors, they have mainly focused on CPU usage, which is considered to be the most important factor in terms of utilization. The algorithms allow equal distribution of resources and prevent migration overhead. Based on this idea, it is very clear that the utilization of resources will be minimized, which allows energy consumption to be reduced. In order to minimize the utilization of resources better, this work will focus on two components, CPU and memory utilization which can be used to better reduce energy consumption and improve performance.

The authors in [24] have also proposed an idea by considering a cooling system and the racks of the network structure that hold the physical servers. They believe these components contribute to energy consumption and performance. This idea goes beyond the normal components in VM migration. They have employed four different algorithms to achieve less energy consumption, quality of service, and optimal placement of VMs. The four different algorithms detect the underutilized and non-underutilized server by selecting the number of VMs from the VMs list by using the *Modified Best Fit Decreasing algorithm* (MBFD), which places VMs in the rack-by-rack algorithm, VMs in the non-underutilized rack algorithm, and migrates VM on underutilized racks consecutively. The whole process uses MIPS based on user's request. Their objective is to take care of consolidation issues which would minimize the number of active racks, which will in turn switch off the unused racks, cooling systems and idle servers after proper migrations have been performed. The research has claimed to improve energy to an average of 14.7% in the entire data center.

Though the authors have considered most of the components that may help in reducing energy consumption and improve performance, it is not very clear how the algorithms will determine the unpredictable tasks sent by cloud users within a 24 hour period as the algorithm has been set to operate on minimum and maximum levels. This research will improve this aspect by focusing on a practical experiment that will estimate the close percentage of utilization that will trigger when the server will become overloaded, under loaded and idle. It will also consider CPU and memory utilization percentage rather than MIPS.

Other research has taken a different approach, which involves memory reusing mechanisms for dynamic VM consolidation [26] in order to reduce the amount of transferred data during migration. They [26] have claimed that the existing live migration techniques take a long time to migrate since they have to transfer the entire image of the memory from one host to another. They have two types of physical hosts; a shared server that handles VMs, which doesn't take on a heavy computation load (close to idle server), and dedicated servers, which take active VMs. In the case that all the VMs happen to be idle, they will be consolidated to the shared servers. If all the VMs happen to be active, they will require higher performance, which means they will need to be transferred to the dedicated servers in order to maintain performance.

As the VM migration happens many times in cloud systems, they claimed there is a very high chance that VMs will migrate to the original host, as has been executed before. For instance, if the VMs happen to move from an idle state to a busy state, there is a need for VMs to be moved from the shared server to a dedicated server to maintain performance. In order to avoid the total transfer of memory from one host to another, an image of memory from one host (i.e. from host Y) can be left and reused when the VMs are transferred back as shown in Fig. 2

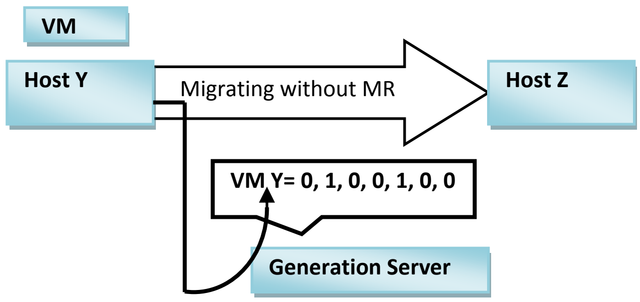

Fig. 2: Migrating VM without memory reuse. Adapted from [26]

Figure 2 shows migrating without memory reuse (MR) during the first transfer; the memory image is stored on the source host (Y). It updates the generation table and the algorithm will know which pages can be reused.

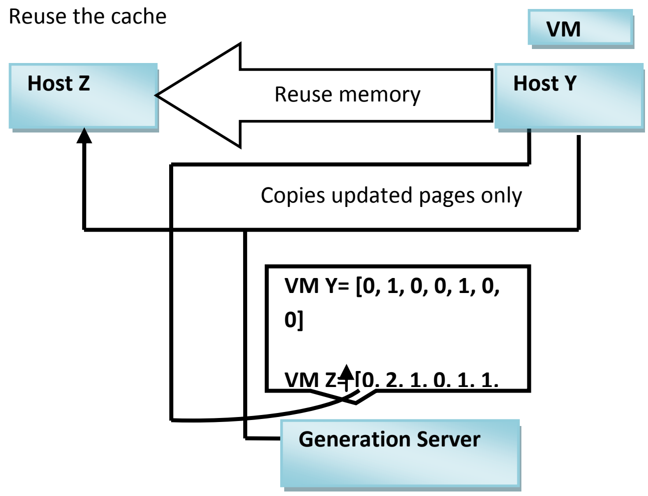

Fig 3: Migrating VM with memory reuse. Adapted from [26]

Figure 3 shows migrating with memory reuse (MR) to host Z, which was migrated the first time. During the migration there is no need for transferring all the pages. The VM will reuse the memory, which was previously cached in host Y.

Since the proposed idea is unique and works in two ways, the system may be able to improve its performance better than a complete image transfer, if the cloud systems are to work as predicted by the authors. This idea may work in an environment where there are a minimal number of hosts and where the original host migration will be easily returned. If comparison is made to the data centers, such as Google and Amazon, it may not always be possible to return to the original location due to the massive availability of servers. In order to be used in a large environment of DCs, this approach will prefer the complete image transfer of memory and CPU utilization from which the delay and better energy consumption might be enhanced.

An algorithm for greener cloud computing has been proposed in [27] by mimicking the behavior of honeybees. The algorithm has made an observation on overloaded, under loaded and idle CPUs. However, the major problem of this algorithm is to find a proper way of managing the idle CPU, the under loaded CPU, and the overloaded CPU. The paper also introduced proper management of overloaded, idle and under loaded CPUs with the aim of turning off the idle CPUs and scheduling the task of the overloaded CPU to the under loaded CPU, for the purpose of saving energy and preventing quality of service violations.

The research in [27] has divided the jobs into two parts; the first part deals with the management of under loaded and overloaded CPUs, in order to reschedule the tasks. The second part deals with the idle CPUs for the purpose of reducing the power consumption.

The research undertaken by [27] seems to have potential, since they have performed practical experiments by using Ubuntu (Linux). The framework proposed is well-organized and linked to the algorithms proposed. The framework is designed in stages and allows cloud consumers to utilize the service based on the limitation and agreement with the cloud providers.

They have used a threshold value of the CPU which is the best way to judge whether the servers are becoming overloaded or under loaded. However they suggested turning off the idle machines which may not be a good idea; if the CPU is turned off for a short time and then rebooted, it tends to consume more electricity, reduce performance and increase the wear-and-tear of server components [28]. The other issue is that if the under loaded CPUs become overloaded by an unexpected request, the question is how an idle server will ensure the quality of service without interruption, since it has to take the process of rebooting into account. As that is the case, this research will propose to put the idle servers into power saving mode to improve performance and reduce energy consumption.

Authors in [29] propose an energy-aware analysis, which is based on RAM utilization (RAM-based host overloading detection), which is quite different from [23] and [30] which are concentrated on the usage of CPU through its utilization and MIPS. Through this research, [29] has claimed to have improved energy consumption by 37.26% and 70% in performance, due to the migration of VMs.

The paper also developed a policy for energy awareness of RAM-consolidated VMs selection. Data center brokers and cloudlets have been created based on the utilization values. VMs, which will be found to have minimal utilization, will be considered to have migrated to other servers.

In order to ensure that the energy and quality of service is maintained, [29] developed an algorithm that will detect the over-utilization and underutilization of hosts, and implement RAM consolidation with local regression techniques, which will evaluate the performance of the host that is combined with the VM selection algorithm. They have redesigned the schedule for the cloudlets in order to get the information of each application. The RAM of each VM has been sampled and added up in each host from the VMs scheduler.

By considering the technique [29] of RAM-based overloading detection, it is clear that the algorithm and the equation suggested will help in improving the performance of the data center. Since the utilization of VMs is calculated by adding the required amount of RAM of the cloudlets, which operates in the VMs and is then divided by the total capacity of RAM in a particular VM, it may not be that easy when applied to the large environment of data centers where millions of users are involved and the task from each user is unpredictable. In addition to that, the technique is able to improve performance and reduce energy consumption as they have considered primary memory alone and living CPU which is considered to be an important component in performance and energy consumption [31]. To overcome this problem, this research will use both components (RAM and CPU) which will also improve performance and reduce energy consumption.

Authors in [15] proposed a novel technique for dynamic consolidation of VMs that are based on the adaptive utilization of the CPU threshold. They have used dynamic threshold utilization, which means the system will be able to adjust based on the behavior of the workload sent by the users. Thus, the system utilization threshold will be based on a history of utilization during the lifetime of the VM operation. This technique is different from other research, which considers the fixed value utilization of the CPU [32],[33]-[34] This novel technique claims to reduce energy consumption and ensures the quality of service to cloud users. The Service Level Agreement (SLA) violation claims to be as low as 1%.

Due to live migration of VMs, the authors have considered the service violation, and have predicted the cost of live migration to be 10%. Based on this claim, they have proposed a dynamic threshold with a heuristic algorithm that minimizes the number of VM migrations.

The proposal of dynamic threshold utilization [35] has provided a different view that will enable a better performance by adjusting the system speed based on the load given by the users. However it may be not easy to limit the VMs migration in an environment of cloud usage such as Amazon or Google, but if migration is limited, there will be performance problems based on the over-utilization of the CPU. The proposed approach will ensure there is a good balance between the overloaded and under loaded servers which will in-turn put idle servers to sleeping mode.

The authors in[36] have come up with a proposal that may help to evaluate server consolidation so as to improve the power efficiency, while ensuring that the QoS is maintained, particularly in large data centers. The model proposed try to address the application workload and considers the effect of energy and QoS while turning on/off the servers and the migration of virtual machine from one data center to another. It mostly concentrates on resource usage (CPU, memory usage and I/O activities). According to the simulation results from the two data centers, which involve 400 concurrent VMs, they claim to have reduced the energy consumption by 50% while ensuring that the QoS is maintained.

The model can improve performance and energy consumption if the allocation of resources in each data center is done systematically. The process is very long, which will in turn consume more energy and absolutely reduce performance. There will be no good reason for terminating the service due to SLA violation after all the processes have failed. The designed agents appears to be fine, but the technique seems to take a longer time, thus making it difficult to match the reality of cloud computing. The technique may work better if the decision of migration is made after the maximum point where the load is not acceptable.

4. EXPERIMENTATION SETUP

In order to test the utilization of the real server, VM consolidation setup have been performed, the setup is based on the ESXi hypervisor on physical machines which will be used to test the overloaded, underloaded, imbalance and idle parameters. The setup includes one high-end server, which is consolidated with 17 VMs. Each VM in the high-end server is installed with a different operating system which consists of servers and clients. The amount of RAM and CPU cores has been allotted to each VM based on their needs. The test bed architecture and set up are shown in Figures 4 and 5.

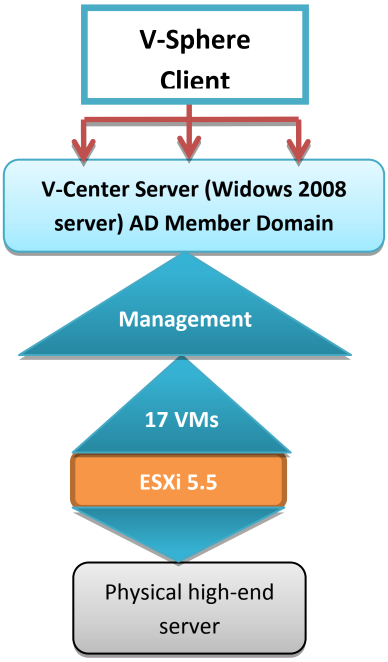

Fig. 4. Test bed architecture

(ESXi 5.5, VSphere Client and V-center)

Figure 4 shows architecture consists of one high-end server which bares 17 VMs working under the layer of ESXi hypervisor. V-Center has been used for the management of all consolidated VMs, and V-Sphere client is used to connect and access the hypervisor remotely.

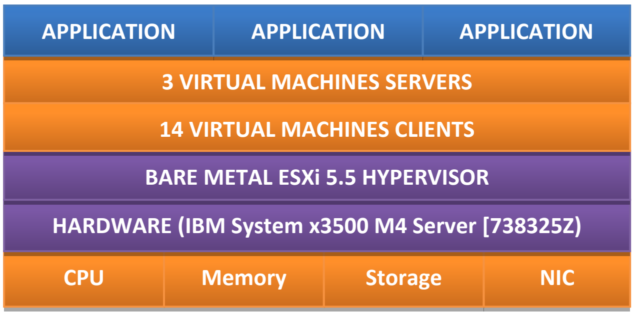

Fig. 5. Cloud test bed setup on high-end server.

Figure 5 shows the system requirements and the setup on a high-end server. The framework has been set up by using a type 1 "bare metal" hypervisor, whereby the hypervisor has been installed on the empty server without any operating system.

5. TESTING

The test will be performed practically by checking the utilization of the CPU and primary memory. Though CPU is considered to be a major contributing factor in energy consumption [23], Primary memory is another contributing factor to degradation of the performance and energy consumption [27][37]. As that is the case, the following equation that includes nodes with CPU and memory utilization is formulated.

$$G_{Cloud} = G_{PMserver} + VMs_{usage}$$

$$G_{PMServer} = G_{CPU_{usage}} + G_{Memory_{usage}}$$

The test will be categorized into four categories in order to investigate all the stages that can be able to tackle energy consumption and performance degradation very efficiently. Each of the below stages will be investigated thoroughly to formulate the accurate parameters that will be used to create a new algorithms for green cloud.

Table.1. Testing Category

Category	Description
Idle Server	This is where the server is completely without any load. It can either receive the load from overloaded or be turned to power saving mode.
Underloaded	This is where the server has minimal tasks to perform; it is between imbalanced. It can either receive the load from overloaded or send the load to imbalanced in order to become idle.
Imbalanced	This is where the server is neither under loaded, nor overloaded or idle. It has a partial workload, and may have some space to accommodate VMs from either overloaded or under loaded servers.
Overloaded	This is where the server becomes nearly fully overloaded based on the utilization of CPU and memory. It can either transfer some of the VM to under loaded, imbalanced or idle.

6. PROPOSED SYSTEM

Though a lot of research is being undertaken in the area of cloud computing to reduce energy consumption and ensure quality of service, comparatively high energy consumption, carbon emission and poor quality of service are still concerns and require more attention to ensure cloud computing becomes greener. The recent research done by [37] still proved high consumption of energy in cloud computing. Energy consumption is not the only issue since consumers of cloud computing are still complaining and doubting the speed of cloud computing.

Moreover there are still more problems on consolidation. Many research and techniques have not specified how consolidation could be better planned and how VMs could be allotted based on different loads sent by different users. The matter is not consolidating VMs on one, two or more physical machines, but consolidation needs critical estimation that will not overload the physical machines and at the same time ensuring the energy is saved. Another important issue is how VMs will be provisioned to take certain tasks from the clients. Also , the problem of utilizing resources to the maximum still persist as the quality of service is not up to the standard of satisfaction, based on the practical aspects, one should consider when and where Physical machines will be overloaded and how to ensure they are not overloaded or under loaded. There are many physical machines as well as virtual machines which are not fully utilized; many services are still distributed separately, leaving lots of idle servers being operated and wasting lots of electricity as well as emitting carbon which pose a threat to the environment.

In order to have Greener Cloud Computing, more active research needs to be done with emphasis on practical experiments. The proposed work involves design and development of a unique techniques/algorithm which will not only reduce energy consumption and lower carbon emission by considering scheduling and machines provisioning, but also ensure better quality of service by taking care of consolidation issues that will satisfy clients and ensure 24/7 service availability. The consolidation issue will be considered practically based on the type and number of VMs used. CPUs utilization and primary memory will need to be allotted to consolidated VMs based on the size, speed of CPU and RAM. A technique based on the consolidation problem will be designed in order to allow data of different capacity to be dynamically allotted to VMs which will be of acceptable CPU and primary memory utilization. After tackling the problem of consolidation, VM provisioning will be another issue; a technique/algorithm will be designed to allow virtual machines not to be used more than overload capacity of total CPU, and primary memory usage. It will also ensure that the server does not stay underloaded, imbalance or idle. The technique will migrate the VMs when the server is overloaded, underloaded and idle. By designing this technique, better energy consumption will be achieved ensure better service to the clients. Another feature of the proposed method is scheduling mechanism by designing a technique that will prevent physical machines and virtual machines being used un-necessarily. The system will be able to monitor unused server (both physical and virtual machines) by putting them in a sleepy mode as well as alerting them to a standby mode in order to be ready to take the incoming tasks.

7. CONCLUSION

In this paper, a detailed review of existing cloud computing algorithms/techniques has been given and the gaps are identified. It is observed that many data centers are still consuming energy, emitting carbon and causing performance degradation to the data centers and end users. Lack of proper planning has resulted into many data center’s servers to be utilized inefficiently. The significance of proper planning and a unique greener cloud computing technique/algorithm is needed to satisfy the needs for energy consumption, carbon emissions and quality of service to the data centers and end users. The proposed unique algorithm will be fully designed and implemented on the physical servers. The test will be performed to check the performance and energy consumptions issues. The achieved results will be benchmarked with the existing algorithms in order to judge the effectiveness in terms of better energy consumptions and better improvement of performance.

8. REFERENCES

[1] S. Thakur, “Server Consolidation Algorithms for Cloud Computing Environment : A Review,” Int. J. Adv. Res. Comput. Sci. Softw. Eng., vol. 3, no. 9, pp. 379–384, 2013.

[2] and B. G. Bedi, Punam, Harmeet Kaur, “Trustworthy Service Provider Selection in Cloud Computing Environment,” in Communication Systems and Network Technologies (CSNT), International Conference on. IEEE, 2012, 2012, pp. 714–719.

[3] C. N. Hoefer and G. Karagiannis, “Taxonomy of cloud computing services,” in 2010 IEEE Globecom Workshops, 2010, pp. 1345–1350.

[4] E. J. Qaisar, “Introduction to cloud computing for developers: Key concepts, the players and their offerings,” in 2012 IEEE TCF Information Technology Professional Conference, 2012, pp. 1–6.

[5] L. Savu, “Cloud Computing Deployment models , delivery models , risks and research challanges,” in 2011 IEEE International Conference on Cloud Computing and Intelligence Systems, 2011, pp. 1–4.

[6] L. Xu, C. Li, L. Li, Y. Liu, Z. Yang, and Y. Liu, “A virtual data center deployment model based on the green cloud computing,” in 2014 IEEE/ACIS 13th International Conference on Computer and Information Science (ICIS), 2014, pp. 235–240.

[7] Urz holzle, “Cloud Computing can use energy efficiently,” 2012. [Online]. Available: http://www.nytimes.com/roomfordebate/2012/09/23/informations-environmental-cost/cloud-computing-can-use-energy-efficiently. [Accessed: 21-Apr-2013].

[8] Emerson Network Power, “Eight reasons IT should measure Energy Cost in Data centers,” 2010. [Online]. Available: http://www.emersonnetworkpower.com/documentation/en-us/brands/avocent/documents/brochures/01/8rsnspwr-ds-en.pdf. [Accessed: 23-Apr-2013].

[9] Energy Star, “Saving energy by using virtualization,” 2013. [Online]. Available: 5. http://www.energystar.gov/index.cfm?c=power_mgt.datacenter_efficiency_virtualization. [Accessed: 11-May-2013].

[10] Oracle, “Strategies for Solving the Datacenter Space, Power, and Cooling Crunch: Sun Server and Storage Optimization Techniques,” 2010. [Online]. Available: http://www.oracle.com/uk/ciocentral/sun-datacenter-space-power-wp-075961.pdf .

[11] B. Yamini and D. Vetri Selvi, “Cloud virtualization: A potential way to reduce global warming,” in Recent Advances in Space Technology Services and Climate Change 2010 (RSTS & CC-2010) IEEE, 2010, pp. 55–57.

[12] and Q. S. Ye, Hanmin, Zihang Song, “Design of Green Data Center Deployment Model Based on Cloud Computing and TIA942 Heat Dissipation Standard Zihang Song , Qianting Sun,” in In Electronics, Computer and Applications, 2014 IEEE Workshop on, 2014, pp. 433–437.

[13] Q. Zhu, J. Zhu, and G. Agrawal, “Power-aware Consolidation of Scientific Workflows in Virtualized Environments,” ReCALL, no. November, pp. 1–12, 2010.

[14] F. Owusu and C. Pattinson, “The current state of understanding of the energy efficiency of cloud computing,” in 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications The, 2012, pp. 1948–1953.

[15] A. B. and R. Buyya, “Adaptive Threshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Data Centers,” Growth (Lakeland), pp. 1–6, 2010.

[16] P. Dalapati and G. Sahoo, "Green Solution for Cloud Computing with Load Balancing and Power Consumption Management," Int. J. Emerg. Technol. Adv. Eng., vol. 3, no. 3, pp. 353–359, 2013.

[17] H. Mi, H. Wang, G. Yin, H. Cai, Q. Zhou, and T. Sun, "Performance problems online detection in cloud computing systems via analyzing request execution paths," in 2011 IEEE/IFIP 41st International Conference on Dependable Systems and Networks Workshops (DSN-W), 2011, pp. 135–139.

[18] S. Acharya and D. A. D'Mello, "Cloud computing architectures and dynamic provisioning mechanisms," in 2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE), 2013, pp. 798–804.

[19] G. K. Sehdev and A. Kumar, "Power Efficient VM Consolidation using Live Migration- A step towards Green Computing," Int. J. Sci. Res., vol. 3, no. 3, pp. 517–523, 2014.

[20] N. Kord and H. Haghighi, "An energy-efficient approach for virtual machine placement in cloud based data centers," in The 5th Conference on Information and Knowledge Technology, 2013, pp. 44–49.

[21] W. Xiaoli and L. Zhanghui, "An energy-aware VMs placement algorithm in Cloud Computing environment," in Intelligent System Design and Engineering Application (ISDEA), 2012 Second International Conference on. IEEE, 2012, pp. 627–630.

[22] A. Beloglazov and R. Buyya, "Energy Efficient Resource Management in Virtualized Cloud Data Centers," pp. 826–831, 2010.

[23] C. Aschberger and F. Halbrainer, "Energy Efficiency in Cloud Computing," pp. 1–16, 2013.

[24] S. Esfandiarpoor, A. Pahlavan, and M. Goudarzi, "Virtual Machine Consolidation for Datacenter Energy Improvement," Comput. Eng., pp. 1–11, 2013.

[25] M. Marzolla, O. Babaoglu, F. Panzieri, M. A. Zamboni, and I.- Bologna, "Server Consolidation in Clouds through Gossiping," in World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE International Symposium on a. IEEE, 2011, pp. 1–6.

[26] S. Akiyama, T. Hirofuchi, R. Takano, and S. Honiden, "MiyakoDori : A Memory Reusing Mechanism for Dynamic VM Consolidation," in Cloud Computing (CLOUD), 2012 IEEE 5th International Conference on. IEEE, 2012, pp. 1–8.

[27] K. Mukherjee, "Green Cloud : An Algorithmic Approach," Int. J. Comput. Appl., vol. 9, no. 9, pp. 1–6, 2010.

[28] Q. W. and N. G. Yiyu Chen, Amitayu Das, Wubi Qin, Anand Sivasubramaniam, "Managing server energy and operational costs in hosting centers," ACM SIGMETRICS, vol. 33, no. 1, pp. 303–314, 2005.

[29] M. A. Bhagyaveni, "VM CONSOLIDATION TECHNIQUES IN CLOUD," J. Theor. Appl. Inf. Technol., vol. 53, no. 2, pp. 267–273, 2013.

[30] A. Murtazaev and S. Oh, "Sercon: Server Consolidation Algorithm using Live Migration of Virtual Machines for Green Computing," IETE Tech. Rev., vol. 28, no. 3, pp. 212–231, 2011.

[31] V. Tiwari, D. Singh, S. Rajgopal, G. Mehta, R. Patel, and F. Baez, "Reducing Power in High-performance Microprocessors," in In Proceedings of the 35th annual Design Automation Conference, 1998, pp. 732–737.

[32] J. L. S. Nath. Alan Roytman, Aman Kansal, Sriram Govindan, "PACMan : Performance Aware Virtual server Machine Consolidation," ICAC, pp. 1–12, 2013.

[33] M. Guazzone, C. Anglano, and M. Canonico, "Energy-Efficient Resource Management for Cloud Computing Infrastructures," Cloud Comput. Technol. Sci. (CloudCom), 2011 IEEE Third Int. Conf., no. CloudCom, pp. 1–11, 2011.

[34] A. Beloglazov, J. Abawajy, and R. Buyya, "Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing," Futur. Gener. Comput. Syst., vol. 28, no. 5, pp. 755–768, 2012.

[35] R. Suchithra, "Heuristic Based Resource Allocation Using Virtual Machine Migration : A Cloud Computing Perspective," Int. Ref. J. Eng. Sci., vol. 2, no. 5, pp. 40–45, 2013.

[36] A. F. Leite, A. Cristina, and M. Alves, "Energy-Aware Multi-agent Server Consolidation in Federated Clouds," Computing, no. i, pp. 1–10, 2013.

[37] and Y. M. Luo, Liang, Wenjun Wu, Dichen Di, Fei Zhang, Yizhou Yan, "A resource scheduling algorithm of cloud computing based on energy efficient optimization methods," in In Green Computing Conference (IGCC), 2012 International, 2012, pp. 1–6.