ISSN (Print) : 0974-6846
Indian Journal of Science and Technology, Vol 9(46), DOI: 10.17485/ijst/2016/v9i46/107108, December 2016 ISSN (Online) : 0974-5645

A Conceptual Framework for Realizing Energy Efficient Resource Allocation in Cloud Data Centre

M. J. Usman[1], A. S. Ismail[1], H. Chizari[1], A. Y. Gital[2] and A. Aliyu[1,3]

[1]Department of Computing Science, Universiti Teknologi Malaysia. 81310, Johor Bahru, Malaysia; umjoda@gmail.com, abdsamad@utm.my, chizari@utm.my, ahmedaliyu8513@gmail.com
[2]Department of Computer Science, Abubakar Tafawa Balewa University Bauchi. 740272 Bauchi Nigeria; asgital@yahoo.com
[3]Department of Computer Mathematics, Bauchi State University Gadau. 751105 Itas-Gadau Bauchi Nigeria

Abstract

Objectives: To present the state of the art on energy efficient frameworks in cloud computing environments **Methods/ Statistical Analysis:** To propose a conceptual framework for energy efficient IaaS (Infrastructure as a Service) of single and multi-cloud data centre. Findings: The approach is based on virtualization and consolidation technique that enables on-demand and dynamic resource allocation while minimizing energy consumption and carbon emission of the data centre with different energy sources. **Applications/Improvements:** The proposed framework unlike the previous approaches support intra and inter-data centre resource provisioning and also deals with dynamic resource allocation of single and multi-cloud data centre.

Keywords: Carbon Emission, Data Center, Energy Efficiency, Network Resource, Resource Allocation, Virtual Machine ,

1. Introduction

Cloud computing recently appeared and become a successful paradigm for providing Information and Communication Technologies (ICT) services and resources based on a concept of pay-per-use basis. The cost effectiveness enabled by the pay-per-use model of the cloud has led to its wide adoption since it saves cloud consumers the cost of purchasing and maintaining both hardware and software resources[1]. Broader acceptance of this paradigm known as cloud computing and virtualization technologies has contributed to the formation of large-scale cloud computing data centers that provide cloud services. These cloud services are being offered in different ways and at a different level of the cloud data centers namely; Infrastructure as a Service (IaaS), Software as a service (SaaS), and Platform as a Service (PaaS)[2, 3].

Cloud services are hosted on data centers, and data are known to consume an excessive amount of energy[4,5]. Consequently, energy consumption by data centers results in high operational cost and it also constitute a significant amount of carbon emission which has anegative effect on the environment. The high energy consumption is not associated with the data center size but rather the inefficient usage of the infrastructure and resource utilization[5]. For these reasons, energy efficiency and resource utilization are becoming increasingly important. Several techniques have been proposed to minimize energy consumption and carbon emission by the data centers, the popular techniques are based on resource allocation, virtualization, or Virtual Machine (VM) migration. Most of these techniques either considered energy consumption of server, Central Processing Unit (CPU), or Memory without considering the networking resources, but the energy consumption of network

* Author for correspondence

resources is not trivial. The few types of research that considered network resource did not take into account the carbon emission as a key parameter for data center selection. Therefore, reducing the energy consumption of processor and network resources need to be considered in one single solution for cloud data center.

There are several means to deal with data center energy consumption. Popular among them is consolidation and virtualization[3]. This technique lets providers of cloud services to create many VMs on a single PM, thereby requiring less physical resources than before, thus improving the utilization of resources. Previous studies mostly focus on improving energy efficiency and resource utilization of Central Processing Unit (CPUs) Memory, storages, and VMs across PM within the data centers. For example, VM placement framework for energy-aware in federated cloud is proposed in[6]. In this framework, the VM placement is possible only by defining constraints stated in the Service Level Agreements (SLA) without considering energy consumption by the network resources of the infrastructure. In[7] proposed a Network framework for Power-Aware using virtual machines re-allocation. The technique provides VM live migration feature that highly supports low energy consumption and improving performance. As can be seen, the frameworks have considered networking resources of the data center infrastructure but still did not take into consideration the carbon emission of the data center base on their energy consumption.

Evolutionary intelligence using Genetic Algorithm has been proposed for VM placement and bi-objective scheduling to minimize energy consumption and violation of SLA. However, the study focuses on energy efficiency and consolidation while ensuring non violation of SLA[8, 9]. There is research that focuses on energy-aware data centers [10, 11]. They proposed frameworks that workout the most energy efficient policies of placing the VM considering resource requirement and selection strategy to minimize energy consumption and SLA. However, other research is concerned with resource allocation and provisioning by putting to sleep the unneeded PM, air conditioning systems, and other IT equipment[12,13]. These frameworks do not consider energy source and carbon emission rate of the data center configurations.

Taking into consideration the environmental factors and quality-of-experience give rise to the needs to developed energy efficient framework known as a green cloud which focuses on environmental impact. Also,some researchers developed a new framework of policies based on User Profile-Aware for Greening the Cloud in order to achieved energy-efficiency and better resource management[14]. In[15] focuses on the green cloud. The authors introduce a green policy for VM manager to consider when triggering VM migration. This operation has played a very inportant role in attaining the goals of green cloud computing. This improves the situation but it is often not adequate due to the dynamicity and heterogeneity of the cloud environment. Consequently, there is a call for energy-efficient resource allocation that will better the services of the infrastructure while maintaining the desired performance level.

In essence, to overcome the limitations of the above-mentioned techniques and to also consider the total energy consumption and carbon emission of the network resources of the data centre infrastructure we are proposing a conceptual framework that integrates the server with network and CPU with network resources using resource allocation strategies and policies. The framework also considers the total energy consumption, off-peak period, the source of energy by all the data center in order to calculate the carbon emission rate. The solution can as well achieve energy savings with all resource constraints (CPU, memory, storage and network) and is expected to be generic, heterogeneous, and adaptive that can be applied to enterprise data centers.

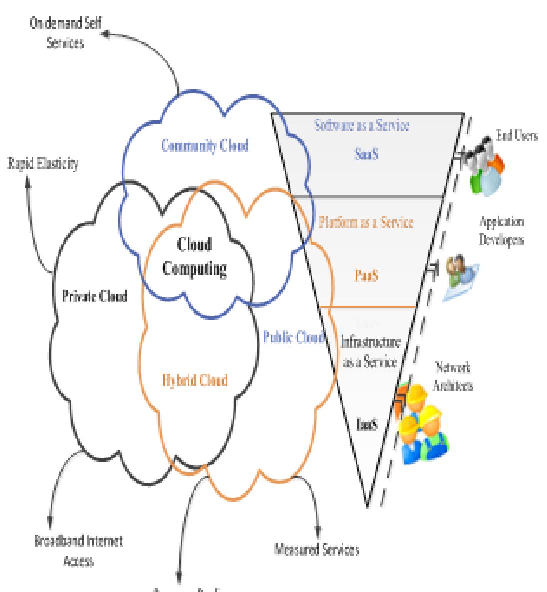
The remaining parts of the paper are organized as follows: Section two discuss the need for cloud computing, types and classification of cloud computing infrastructure, and its service models. The challenges of cloud data center resource provider are discussed in section three. Energy efficiency in cloud computing data center resources is also presented. Section four presents the proposed framework followed by the components of the framework including some approaches. Section five concludes the paper and discusses the future works.

2. The Need for Cloud Computing

Cloud computing is essential nowadays. The days are not that far when people, organizations, research institute, and any functional agency will be passing on cloud computing to execute the business and non-business related tasks. Therefore, researchers proceed to scalable cloud data centers that allowed for cloud computing resources to be leverage (or, possibly more accurately rented) on request by the user and in a scalable manner through the internet. Approximately in 2007, the term cloud computing made an inroad into computing and people started using it to ponder on the use of the internet technology in relation to that available services[16]. A public organization such as research institutes, military, healthcare system, and private organization are speeding to embrace cloud computing technology to work out their demanding IT related problems arising from the digital information era. According to[17] there are three important components that contributed to the rise and concerns in cloud computing environment. They are:

- The decrease in hardware cost and increment in computing storage power and capacity, and the advent of modern supercomputers and multi-core architecture comprising of many cores if not in thousands.
- Recent development in big data and the exponential rise of scientific experiment, simulation, E-transactions and online-storage.
- The general acceptance and widespread use of computing services and applications Web 2.0.

All of above mentioned attributes have contributed a lot to the rise and move of cloud computing, by acquiring the power to access it from any nook and crannies of the world, is nothing less than a dream[17]. Also according to Gartner report[18] the movement to public cloud computing services is among the most raging issues in ICT industry. As further detailed by the report, by 2015, 50% of Global i.e about 1000 enterprises demands will have to depend on outsourcing services of the cloud computing as one of their major sources of revenue, and by 2016, almost all the Global 2000 industries across the nations will be using some level of public cloud services which we are already witnessing now. Figure1 shows the classification and functional aspect of the cloud computing, which includes the services, carriers, models, as well the characteristic that uniquely identify it as the way to go.



The majority of frequent proficiency of presenting cloud service models could be segmented into SaaS, PaaS, and IaaS. In this kind of model, users usually access the services on the basis of their request without knowledge to where the services are being hosted[19]. It is significant to mention that, as shown in Figure 1, the SaaS as a service delivery is built on PaaS, and the latter on IaaS. Thus, this is not a leaving out approach to categorization, but rather it is relating to the level of services provided by the cloud model. IaaS offers computing resources such as storage, networks, and processing which establish a set up for PaaS services, and PaaS offers tools and services for application development and deployment. It is able of performing precisely the same for SaaS services. SaaS offers complete application services like calendars, customer relation management, office suites e.tc, it can be hosted on top of the previous service model and all of them will form on top of IaaS which is the overall architecture of the cloud computing data centers.

Users still have the right to make a decision between using private clouds, which are DCs (data centers)design for the internal used of a certain institution and or business organization, and public clouds, which are often free for the public to use over the Internet. The private and on-premises public cloud are combined together on a single cloud with orchestration and automation between the two to form a hybrid cloud for the organization that needs extra services and it's particularly valuable for dynamic or highly changeable workloads. The community cloud is usually controlled and operated by a group of institutions or organization that share common interest and objective. Since the resources are shared by specific communities there is no issue of resource management whether hosted externally or internally[19].

3. Challenges of Cloud Provider and Energy Efficiency

The providers of the cloud datacenter services are responsible for allocating resources to service providers to maximize the total income, while minimizing operational expenditures. The fundamental functions in cloud computing environment include service users and the service providers. The cloud computing service users need a secure point of connection to access low-cost services that are elastic and comfortable to use anywhere and anytime.

Table 1. Cloud perspective

Cloud Provider Perspective	Cloud User Perspective	End User Perspective
Deploy platform over which the available resources in the IaaS will be provision for users on demand request.	Enable users have access to applications from anywhere anytime.	Generates the workloads.
Provide resources such as server, storage, memory, CPU, networks in form of virtualization	A system, which is virtualized and automated.	Requests for applications hostedon virtualized infrastructure.
Responsible for provisioning of space, power & cooling.	Responsible for deploying application and data on the platform.	Applications are often run on specialized software platforms.
Responsible for provisioning of load balancing e.t.c.	Maintenance by the provider.	Maintenance by the provider.

3.1 Cloud Provider Perspective

The service provider usually operates the requested applications or the service that the user is needed[20]. Table 1 below outlines the perspective of the service providers and the customers of the IaaS data center from a cloud resource management perspective.

The operational expenditures in data center environments include costs for server, network maintenance, energy (power and cooling) cost, as well as maintenance cost.

3.2 Cloud Provider Challenges

Cloud providers are facing some management challenges that are highlighted below, in order to maximize leasing of resources and to minimize the consumption of energy in the cloud data centers.

- **Workload Scheduling:** As mentioned previously, cloud computing data centers usually comprise of physical machines with resources that are heterogeneous in capacities and its functioning features. Simultaneously, workloads in cloud environment indicate important diversities in terms of resource requirements, antecedence, performance objectives, and characteristics. Therefore, it turns out to be a disputing issue to decide the optimum assignment of resources to applications that are ready to satisfy application performance objectives while maximizing the total utilization of the physical resources [21].
- **Energy Efficiency:** Cloud computing data centers usually consumed a substantial quantity of energy to carry out its functions. Meanwhile, cloud providers need to make a strong motivation to improve the efficiency of the data center energy consumption. One of the ultimate objectives of energy management is to make data centers energy-proportional, meaning that the energy consumption should be proportional to the actual resource usage at the data center. However, commodity machines today are far from being energy proportional. It has been reported that, physical machine or servers when idle consume 50% of energy[22,23]. Therefore, the most effective way to save energy is to turn the machines off when they are not in use. Various techniques, such as server consolidation, have been shown to be effective for minimizing the number of used machines.
- **Pricing:** Pricing is another important issue not only because it directly affects the revenue of the cloud provider, but also because it can influence the allocation strategy of the service providers, which will eventually affect the effectiveness of a resource allocation scheme. Currently, most of the data center service providers such as Amazon EC2 and Rackspace lease virtual resources at a fixed price[28]. However, recent work suggested that the charging of flat rate system can contribute to ineffective effects, due to the mismatch between resource availability and demand fluctuation[24]. A common resolution to this problem is to adapt the price, based on the demand and supply using dynamic pricing schemes such as auctions.
- **Resource Allocation:** Resource allocation (RA) has a substantial effect on cloud data centers, especially in pay-per-use deployments scenario[25]. Allocating the right resource that will perform the required task with minimum cost and execution time is the major issue. Once the needed types of resources are decided, then

instantly these resources are allocated to execute the task. Resource determination and allocation for each task is managed by task modules. One of the main ideas in resource provisioning is to make resources available based on applications requirement in a way that reduces energy wastages, costs, there by utilizing the available resource.

- **Performance:** In cloud computing, performance is one of the keys for measuring efficiency. For data center to increase its performance and utilization of resources. They need to employ consolidation and virtualization techniques unlike the traditional data centers, as a way to provide support to simultaneously running applications on a large number of PMs with different scheduling policies. In the multi-cloud scenario, where we have more than one provider, performance is of the additional concern because of its significance and may differ significantly[26]. Preserving data centers performance that is designed by integrating resources of different types is a challenge with high complexity.

3.3 Energy Efficiency in Cloud Data Center

Cloud infrastructures constitute networked of thousands of PMs with their associated storages, processors, Power Distribution Units (PDUs), air conditioning and other subcomponents. Because of the equipments sizes and number, data centers consume high energy and also let out high carbon. Report on g data centers computing in 2007by US Environmental and Protection Agency (EPA), the energy consumed by cloud data is almost 1.5% of US total energy[27]. The higher the energy consumption the higher the carbon emission which were translated into 80-116[27] metric mega tons each year. Figure 2 identifies and lists the equipment which is typically used in data centers. Developing cloud model that is energy efficient does not show that the hosts or PMs and their subcomponents within the infrastructure are also energy efficient. There are some existing components as shown in Figure 2 that needs to be considered in order to get the overall energy consumption of the infrastructure.

4. The Conceptual Framework

The proposed framework presented will integrate cloud data center resources such as processor and network, as well as server and network that enable on-demand, and dynamic resource allocation while reducing the energy consumption of the data center having different energy sources and carbon emission rates. It has been observed that research in the area of cloud data center does not explore the integration of IaaS infrastructure on a single and multi-cloud data center. The approach is virtual based and consolidation of the IaaS. This approach permits one to overcome power clumsy by accommodating multiple VMs on a single physical host and by performing live migrations to optimize the utilization of the available resources. The components are defined by the current data center demand for provisioning, optimization and the location of the data center where we can migrate our resources when the need arise. The constraints such as energy availability, environmental impact, off-peak time, and automation are the problems associated with the demand as illustrated in Figure 3.

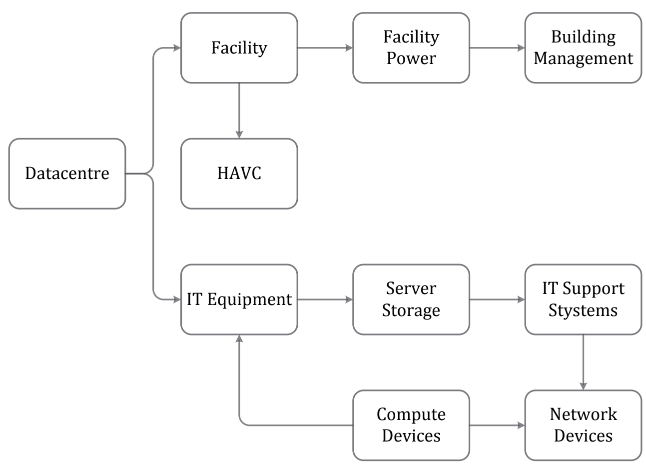


Figure 2. Subcomponents of data center infrastructure.

Efficient resource allocation and energy consumption are needed by both cloud provider and users for improving DCiE (Data Center infrastructure Effectiveness) and timely delivery of jobs. As seen in Figure 3 above, the framework is divided into three layers; Hypervisor layer, Infrastructure layer that consists of optimization module, and user layer which are built on each other to achieve the desired objective. dataset. A training set is constructed using the standard health data as per the listed search words. Then the raw (test) data are compared with the training data and classified into groups. This data can be utilized for continuous advocacy monitoring.

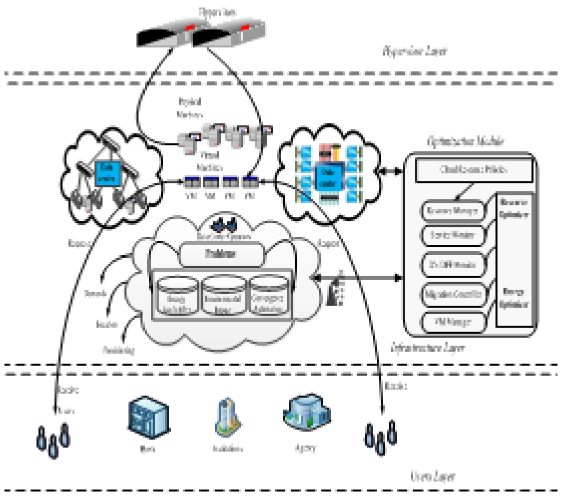


Figure 3. Proposed framework for energy-efficient resource allocation.

- **Hypervisor Layer:** This hypervisor technology is based on a software layer that manages the resources of a physical machine or servers in cloud data center. This will provide support to the needed services for the VMs to run. Rather than to provide direct access to the available resource of the infrastructure layer, all VMs request by users are treated at this layer that control virtual resources. In an environment with hypervisor layer, cloud services can be encapsulated since the underlying resources are allowed to run on top of the hypervisor their by allowing efficient resource allocation and monitoring of their usage by the cloud service provider. Given these points, we can agree that with this layer where clients or users received virtual resource can increase the performance of the data center infrastructure as well optimize resource utilization.
- **Infrastructure Layer:** Provision of resources including physical and logical (servers, processors, networks, and virtual instances)to users is achieved by utilizing hardware virtualization technique. Users and or consumers of cloud services, acquire resources from cloud service provider based on IaaS model. That will save them from maintaining an infrastructure with committed logical and physical resources. Although, there are various infrastructure resource service providers available in the market such as Amazon Elastic 6 Compute Cloud (EC2)[8] and GoGrid[29].

The optimization module where every VM has its own characteristics, features, and functionalities upon user requirement and or specification. The total energy consumed by PMs depends upon the usage of resources. Thus, the total carbon emission of the data center depends on the energy consumption by PM and network components and can be handled by resource and the energy optimizer. The remaining components in the layer perform the function of migration and VMs allocation, monitor resource utilization to decide ON/OFF resource based on the cloud policy been adopted.

- **User layer:** End-users in cloud computing paradigm are considered as the most significant because they are the consumers of the services provided by the IaaS and usually have no concern on where these services are hosted. However, they expect authentic services, that are been delivered on time. The connection points or interfaces, collaborative support, information about their services, etc. should be very simple and easy to use. The distribution of services and resources across the data centers will usually depend on the constraints define in the SLA metrics for the IaaS. These metrics include: service availability, scheduling policy, task, CPU capacity/ Storage, size of the RAM, and execution time for request completion by the user.

5. Conclusion

Servers and their associated CPU with network components are the highest consumers of energy in the cloud data center which makes them a real target for optimization. However, cloud service providers rely on large and energy consuming data centers to support the scalability and elasticity demand by their clients. Energy-efficient resource allocation remains a big issue for cloud providers as increasing demand, provisioning, and high-performance expectations which are also challenging in the cloud model due to the dynamism of its components and flexibility of services supported by virtualization. In this paper, the functional component for achieving energy-efficient resource allocation for single and multi-cloud data center are presented and, based on their demand the most significant standards and or measures for realizing energy-efficient resource allocation are defined as four different layers in order to help data center operators maintain a reliable and performance guaranteed services to the cloud users. Finally, we present three scenarios; processor with network resources; a server with network resources; and their integration in a multi-cloud, which can be applied to the cloud data center environment when clients demand more resources from their cloud provider. As future work, we will design and implement an algorithm using swarm intelligence in Cloud Sim environment. The simulation process will serve as a test bed of our framework. The result of this work will be the basic tool to be used with trust management systems for cloud computing to help data center providers achieve energy-efficient resource allocation and low carbon emission with desired performance level.

6. References

1. Buyya R, Yeo CS,Venugopal S. Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering It Services as Computing Utilities. Paper presented at CCGRID 2009. Proceeding of the 9th IEEE/ACM International Symposium on Cluster Computing and Grid, Shanghai, China. IEEE;IEEE; 2009 May 18-21; p. 5–13.
2. Foster I, Zhao Y, Raicu I, Lu S. Cloud computing and grid computing 360-degree compared. . Proceeding of the 2008 Grid Environments Workshop, Austin, TX USA. IEEE; 2008, Nov 12-16, p. 1–10.
3. Beloglazov A, Abawajy J, Buyya R. Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. Future generation computer systems. 2012 May 31; 28(5): 755–68.
4. Aslekar A, Damle P. Improving efficiency of data centres in India: A review. Indian Journal of Science and Technology. 2015 Feb 1; 4(8):44–9.
5. Beloglazov A, Buyya R. Energy efficient resource management in virtualized cloud data centers. 10th IEEE/ACM international conference on cluster, cloud and grid computing. Melbourne, Australia. IEEE; 2010, pp. 826-831.
6. Dupont C, Schulze T, Giuliani G, Somov A, Hermenier F. An energy aware framework for virtual machine placement in cloud federated data centres. 3rd International Conference on Future Energy Systems: Where Energy, Computing and Communication Meet (e-Energy),Madrid. Madrid. 2012, p. 1–10.
7. Alhiyari S, El-Mousa A. A Network and Power Aware framework for data centers using virtual machines re-allocation. Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Jordan, Amman. IEEE:2015, p. 1–6.
8. Theja PR, Babu SK. An evolutionary computing based energy efficient VM consolidation scheme for optimal resource utilization and QoS assurance. Indian Journal of Science and Technology. 2015 Oct 21; 8(26):1–11.
9. Beegom AA, Rajasree MS. Genetic Algorithm Framework for Bi-objective Task Scheduling in Cloud Computing Systems. 11th International Conference on Distributed Computing and Internet Technology, ICDCIT, Bhubaneswar, India. February 2015, p. 5–8.
10. Zhang Z, Hsu CC, Chang M. Cool Cloud: A Practical Dynamic Virtual Machine Placement Framework for Energy Aware Data Centers. 8th International Conference on Cloud Computing.Jun, 2015, p.758–65.
11. Zhou Z, Hu Z, Li K. Virtual Machine Placement Algorithm for Both Energy-Awareness and SLA Violation Reduction in Cloud Data Centers. Scientific Programming. 2016 Mar 31; 2016:1–11.
12. Kumar BS, Parthiban L. An Energy Efficient Data Centre Selection Framework for Virtualized Cloud Computing Environment. Indian Journal of Science and Technology. 2015 Dec 12; 8(35):1–6.
13. Dabbagh M, Hamdaoui B, Guizani M, Rayes A. Energy-efficient resource allocation and provisioning framework for cloud data centers. IEEE Transactions on Network and Service Management. 2015 Sep; 12(3):377–91.
14. Alhaddadin F, Liu W, Gutiérrez JA. A User Profile-Aware Policy-Based Management Framework for Greening the Cloud. 4th International Conference on Big Data and Cloud Computing (BdCloud). Dec 2014, p. 682–7.
15. Khajehei K. Green Cloud and Virtual Machines Migration Challenges. Indian Journal of Science and Technology. 2016 Feb 9; 9(5):1–8.
16. Yeluri R, Castro-Leon E. Building the Infrastructure for Cloud Security: A Solutions View. Apress; 2014 Mar 27.
17. Prasanth A, Bajpei M, Shrivastava V, Mishra RG. Cloud computing: A survey of associated services. Book chapter of cloud computing: Reviews, surveys, tools, techniques and applications-an open-access eBook published by HCTL open. 2015, 1–15.
18. Lin A, Chen NC. Cloud computing as an innovation: Perception, attitude, and adoption. International Journal of Information Management. 2012 Dec 31; 32(6):533–40.
19. Marshall P, Keahey K, Freeman T. Improving utilization of infrastructure clouds. Proceeding of the 11th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. Melbourne, Australia. IEEE; 2011 May 23-26. p.205–14.
20. Madni SH, Latiff MS, Coulibaly Y. Resource scheduling for infrastructure as a service (IaaS) in cloud computing: Challenges and opportunities. Journal of Network and Computer Applications. 2016 Jun 30; 68(2016):173–200.
21. Younge AJ, Von Laszewski G, Wang L, Lopez-Alarcon S, Carithers W. Efficient resource management for cloud computing environments. International Green Computing Conference. Chicago, IL. 2010,p. 357–64.

22. Younge AJ, Von Laszewski G, Wang L, Lopez-Alarcon S, Carithers W. Efficient resource management for cloud computing environments. Proceeding of the 10th Green Computing Conference, Chicago, Illinois, USA.IEEE; 2010, p. 357–64.
23. Zhang Q, Zhu Q, Boutaba R. Dynamic resource allocation for spot markets in cloud computing environments. Proceeding of the 4th IEEE International Conference on Utility and Cloud Computing, Melbourne, Australia. IEEE; 2011, p. 178–85.
24. Lee YC, Zomaya AY. Energy efficient utilization of resources in cloud computing systems. The Journal of Supercomputing. 2012 May 1; 60(2):268–80.
25. Tickoo O, Iyer R, Illikkal R, Newell D. Modeling virtual machine performance: challenges and approaches. ACM SIGMETRICS Performance Evaluation Review. 2010 Jan 21; 37(3):55–60.
26. Gough C, Steiner I, Saunders W. “Why Data Center Efficiency Matters,” in Energy Efficient Servers: Blueprints for Data Center Optimization. Berkeley, CA: Apress, 2015: p. 1–20.
27. Barroso LA, Hölzle U. The case for energy-proportional computing. IEEE; 2007; 40(12):33–7.
28. Amazon.com, Inc. Amazon Elastic Compute Cloud. http://aws.amazon.Com/ec2/. Date accessed: 04/ 2016.
29. GoGrid. GoGrid Cloud. 2016 April. Available from: http://www.gogrid.com.