IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-6, Issue-9, Sept. 2018 E-ISSN: 2347-2693

A Taxonomy and Survey of Energy Efficient Resource Allocation Schemes for Cloud Datacenter

Satveer[*], Mahendra Singh[2],

[1*]Department of Computer Science, Gurukul Kangri Vishwavidyalaya, Haridwar, India
[2] Department of Computer Science, Gurukul Kangri Vishwavidyalaya, Haridwar, India

Available online at: www.ijcseonline.org

Accepted: 26/Sept/2018, Published: 30/Sept/2018

Abstract— The cloud computing is gaining the popularity rapidly due to its esteem services benefits and high computational demand of social, business, web and scientific applications. The cloud datacenters across the world are consuming high volume of energy thus affecting the environment also. The resource allocation in cloud is a key factor to achieve energy efficiency. In this paper, we reviewed the energy concept and different kinds of resource allocation schemes being used in cloud datacenters and went on to derive a taxonomical classification of these strategies based upon various metrics.

Keywords— Energy efficiency, QoS, Virtual Machine, Datacenters, Cloud Service Provider, Resource Allocation**.**

I. INTRODUCTION

Cloud computing has been speculated as the next-generation forthcoming computing model, due to its compliment benefits in terms of on-demand service, ubiquitous network access, site free resource pooling, and transposition of the risk [1]. Cloud technology shifts the computational and storage resources from the network core to cloud, so that users can access the resources from anywhere on demand and cloud service provider charge their users on the basis of pay-as-you-go model [2]. As the cloud gaining popularity continuously due to its feasibility, CSPs are establishing power hungry mega datacenters across the world to provide different types of services i.e. computation, storage, and networking [3]. These DCs consist of different resource capacities and performance levels servers or called Physical Machines of various generations and configurations. However, DCs are consuming enormous amounts of energy that surge total operational cost and carbon footprint to execute the user workload [4]. A smart report of Global sustainability Initiative (GeSI) stated that, the ICT DCs will be amenable for 18% CO_2 ejection and carbon emission will be 1430 million metric tons [5]. The figure 1 demonstrates the CO_2 ejection by worldwide DCs nominal by Standford University, MCKinsey study and Gartner research [6]. Mostly the hardware devices consume their maximum power level when they are idle, due to resources underutilization. According to [7] servers use nearly 30% of their peak power while sitting idle 70% of the time. The maximum power in DCs are consumed by hardware infrastructure including servers, storage, networking devices and cooling systems as well computing resources executing the user workload and which further gets transformed into heat emission. This extremely high temperature in DC invokes the problems such as system reliability, availability and reduces the life time of the system. Virtualization helps to reduce energy consumption through efficient resource utilization in cloud datacenters [8]. The reason for this enormously tall expenditure of power electricity is not just lies in the amount of computation tools employed, or the energy inaptness of hardware, but also due the inefficient allocation of resources. The inefficient resource allocation of the resources not only reduce Total Cost of Ownership (TCO), but also influence the environment negatively [9]. In present day, there are many penetrating key intentions in front of CSPs such as meeting performance constraints under uncertainties, dynamic scalability, standardization, fault-tolerance and reducing energy cost [10]. The major objective of the CSPs is to model cost effective and power efficient management for itself as well as their end users. Minimizing the carbon footprint produced has appeared as one of the crucial research context both in industry and academia persons. Therefore, CSPs need to renovate the resource allocations strategies for achieving the best trade-off between energy efficiency and QoS.

The present paper addresses the problem of energy efficiency with resource allocation and proposes a detailed taxonomy of energy efficient resource allocation techniques. Rest of the paper is organized as follows: The relationship between energy and power and the concepts of energy efficiency and resource allocation are discussed further in this section. The section 2 correlates the resource allocation with energy efficiency in cloud DCs. The section 3 articulates the taxonomy classification of energy efficient resource allocation schemes and the last section contains the conclusion.

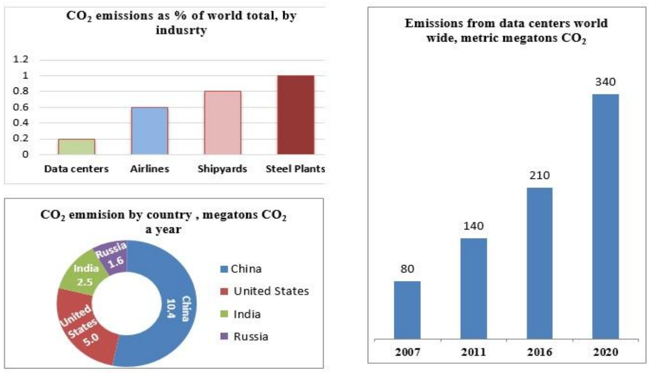

Fig. 1 CO_2 Emission from Datacenters Worldwide.

A. Relationship between Energy and Power

In the context of energy efficiency for cloud datacenter it is essential to realize basic relationship between the energy and power. The energy and power are nearby connected to each other. Electric current is the flow of the electrons in electric conductor, which is measured in ampere. Current defines amount of electrons move aside by an electric circuit per second. One ampere can be defined as the degree of the current generated by potency of one volt acting through a one ohm resistance. The amount of work performed by a system is called as the power. Power is the rate or frequency or rate at which the system executes works and it is measured in Watts (W). While on other side, the total unit of work competed over duration of time is known as the energy [7]. The energy is measured in watts per hour (W/h). The rate of energy per unit time and energy consumed per hour are measured by the power and energy respectively. The symbolic relationship of power and energy can be formulated as follows:

$$P = \frac{W}{T} ; E = E = PT; W = E$$

Where T is a period of time, W denotes the overall work completed under that epoch of time. In addition, the symbolic constants P and E represent the power and energy respectively. For reducing the energy consumption in cloud datacenter it is necessary to state the distinction between power and energy. Because by always decreasing the power consumption it is not possible to lowering the energy consumption. The reason behind this is that, by decreasing the CPU epochs for an executing program ever does not always lead to reduction in adequate power consumption. Nevertheless, in this circumstance the executing process may postulate large amount of time to accomplish its execution on the same unit of energy consumption [7].

B. Energy Efficiency

The major seed of energy sources for managing energy crisis are such as fossil fuels, coal, water, wind, oil, natural gas, and other renewable and non-renewable sources. India also generating their energy from coal based fuel [11]. However, energy demand can partially be fulfilled from non-renewable (coal) sources, due to limited number of reserves and deficit. A report of World Energy Outlook in 2013 stated that the renewable energy sources will be the second largest energy sources in the world, due to regular growth of hydropower and bio-energy [12]. However, theses renewal and nonrenewal energy sources are spot and weather dependent and have their respective impactions [13]. A modern scalable protraction for energy efficient cloud infrastructure is desired that can result in the deficiency in the Greenhouse Gases [14][15][16][17]. The enhancement in ICT resource quantity and density has straight influence on the end user expenses and the TCO [18] [19]. To optimize the operation of a cloud DCs, it is crucial to consider minimizing energy consumption by including compute and cooling components. Today, there are four well conversant perspective to construct the energy efficient DCs: (1) minimizing the power radiate in ICT infrastructure by the well-tuned decision four resource allocation (2) Securing the availability of infrastructure to decrease the necessity of the resource replacement (3) Enhance the resource utilization by dynamically allocating resources (4) Reducing the self-management and flexibility as the cost is shared among cloud providers [20]. Generally in DCs two well-known strategies: (a) resource virtualization and (b) sleep scheduling that helped to enhance the energy efficiency. One of the global campaigning organization Greenpeace surveyed and examines reality check of top five CSPs and reported research facts [21]. According to their report it was uncovered that only few CSPs are interested toward moving the usage of clean and renewable energy. The CSPs must rise with use of more clean and renewable powers. The power efficiency of a DC can be measured by accounting the two parameters: (1) fuel (coal) intensity; (2) clean energy index, the fuel intensity and clean energy index calculated and compared with the greatest power requirement of the DC. The fuel intensity is the imminent scale of electricity originated by coal for DC functions and the measurement of the clean energy index is based on the amount of electricity required and the renewable power used to actuate the DC. At present scenario every CSPs are trying to find the relation between power consumption by cloud ICT resources and their CO_2 emission rate.

C. Resource Allocation Problem

The resource allocation can be defined as the sequence of resource discovery, selection, provisioning, task scheduling, and management of resources. The schematic procedure of how much, what, when and where to allocate the available resources to the cloud consumer application are also determined in the resource allocation [22]. In cloud paradigm, resource allocation is one of crucial challenge because lots of clients are generating fluctuating demanding workload at all time. The resources in cloud are available in the form of global resource pool, where the request for resource cannot be generated directly, yet can be achieved by using the Simple Object Access Protocol (SOAP)/Restful Web APIs. The SOAP/Restful maps the user request to the demanding ICT resources (CPU, blob Storage, Network bandwidth, and system cache). Cloud is potentially viable to boost on-demand resource allocation because it manages the resources into the form of abundance of resource. However, this type of management of resources leads to non-optimal resource allocation in cloud. Numerous cloud consumers generate a large no of unpredictable and changing request along with their QoS requirements simultaneously at any time with specified amount of resources and type that they need [23]. Generally, cloud computing composed of two entities: CSPs and cloud user and offers the multitenant environment, where multiple clients generate multiple requests. The Cloud User generates a large volume of unpredictable and changing request along with different QoS requirements simultaneously at any time with specified amount of resources that they need. While on another side after receiving a CUs workloads, the CSPs analysis the matching of task demand to extant computation resources and determine its presumption, whether consumer task demand can be served on available computation resources along with QoS needs or not. Subsequently after getting the resources the CUs run their applications. When the consumers accomplish their task they lease back resources to the cloud provider. The basic flow of cloud resource allocation is shown in figure 2. Another delicious fact between both parties is that they have their own specific target of monitory benefits. Usually, CSPs wants to gain as many as possible monetary with using least amount of infrastructure or in other words, they willing to maximize TCO throughput by initiating maximum number of VMs with using minimum PMs, but can result in performance degradation. While on another side, user wish to minimize the economic expenditure by expecting that the task can be completed on time. In cloud both these parties are very greedy in nature and leads to impairment for resource allocation decision making, because cloud consumers (do not expose the nature of task, source code, data set) and cloud provider (do not expose numbers and types of PM) do not share information to each other. Thus the CUs never exposes their task resource demands because they do not know what exactly available. Similarly cloud providers cannot allocate resource most suitable for user's workload, because they know the few details about the user. This confidentiality of both the parties influences the resource allocation [24]. The major objective of Cloud Service Provider regarding the resource allocation is the analysis of the workload's characteristics according to available resources, and reduces power consumption by energy efficient resource allocation. Optimizing the QoS parameters scales the proficiency of resource allocation rather than optimizing the type of resource distributed to the end-clients. The problem of resource allocation must be based on such criteria as Quality of Service (QoS), service needs, time, budget, energy efficiency, and effective utilization of resources [25]. The CSPs allocates DC resources in dynamic fashion for their users and charge them pay per used based and granted reliability oriented services.

Resource allocation problem is absolutely sophisticated due to unpredictability of workload and requires certain assumptions considering: (1) set of workflows task for resource needs (2) set of functioning PMs,(3)Workload consolidation according to SLA,(4) reduction in total operational cost by reducing power consumption and resource cost. Hence, the problem domain of resource allocation involves reasonable resource provisioning and need renovation according to the energy efficiency and QoS balance.

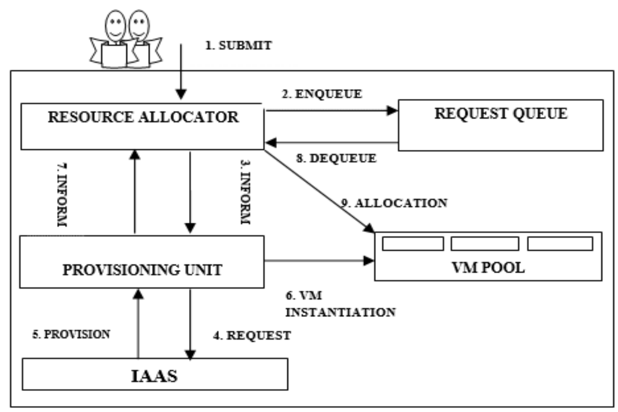

Fig.2. Elementary Cloud Resource Allocation [26].

II. RESOURCE ALLOCATION WITH ENERGY EFFICIENCY

The intention of energy inefficiency in CDCs across its infrastructure is crucial and has accepted broad identification in cloud. The energy efficient resource allocation problem in cloud can be defined to be as "the problem of selecting, deploying, allocating, runtime adjustment of the instant available resources in a way the requested services can obtain their QoS in inexpensive way, so that the resource utilization can be increased as well as the financial and environmental cost can be reduced"[27]. Energy efficient means that the compatible resource allocation must made be for appropriate user workload at appropriate amount of energy cost. In addition, energy efficient resource allocation not only maximizes the throughput, but also reduces the amount of resources that helps in saving energy. Some key points that are very helpful for making energy efficient resource allocation decision: (a) system monitoring to monitor the performance of resources, (b) well-tuning and tracing the QoS needs, (c) tracking the user's task request (d) determination of the price according to the resource usage, (e) Workload consolidation according to SLA (f) keeping the resource usage history, so that the resource allocation can be made intelligently. A modern scalable mechanism for energy efficient cloud infrastructure is desired that can result in the deficiency of CO_2 emission. At present energy efficient allocation of resources in cloud computing paradigm has been an affinity for researcher [28].

III. Taxonomy of Energy Efficient Resource Allocation Schemes

The cloud computing resource allocation strategies are designed by employing the various techniques to produce the best trade-off between energy and SLA. Although, one interesting key point which we have noticed in literature is that the resource allocation is also recognized with another signatures such as dynamic resource allocation, resource management and Virtual Machine Consolidation (VMC). The dynamic VMC is the most widely used approach for saving energy with low SLA violation and power preserving is obtained by bringing the idle or under loaded hosts or servers into power saving modes with the use of VM consolidation. We have designed a taxonomy of most energy efficient and prominent resource allocation schemes that have been formed to gain energy efficiency with low SLA violation for cloud DCs. This section explains a thematic taxonomy based on various metrics such as resource adaptability, optimization objectives, allocation methods, optimization strategies, allocation operation, target resource type, target system and resource level. The taxonomy has been presented in fig. 3.

A. Resource Adaptabilities

The resource adaptability attribute indicates the degree to which an energy aware resource allocation is potentially viable to accommodate to dynamic circumstances also called uncertain condition. Based on the nature of techniques the allocation strategies can be divided into three categories: (1) predictive, (2) reactive, (3) hybrid.

In order to adjust the resource allocation, predictive resource allocation technique can anticipate and seize relationship among service QoS goal, energy efficiency, present allocated resources and workload patterns of the users by using past knowledge driven machine learning techniques. This past knowledge is earned through the monitoring activity. In prediction technique resource allocation depends on the feedback outputs of the machine learning techniques such as neural network genetic algorithms, reinforcement learning [29][30]. The reactive resource allocation adapts availability of the resource, application requirements and analyses long term resource allocation. The reactive mechanism is very attractive because of two reasons no previous resource performance data and workload requirements. Reactive approach responds fast to changes in workload demands but has few importance because of some issues like; lack of forecasting, instability and high provisioning prize [31].The workload behaviour is predicted by the workload predication models by accounting the resource requirements of applications. While the performance of CPU, storage, and network is forecast by the resource performance model. The predication resource allocation techniques depend on resource performance data and particular distributions. The prediction mechanism is costly in terms of storage and processing time complexity. Hybrid resource allocation combines predictive with reactive allocation techniques and accomplishes substantial improvements in: meeting SLAs, conserving energy, and reducing provisioning costs. Hybrid allocation approaches out performs predictive and reactive resource allocation strategies when performance, power consumption, and number of changes in resource allocation are considered [32].

B. Optimization Objectives

Each resource allocation strategy has its own different optimization objectives: Energy efficiency yield to monetary profits by reducing the power consumption. Efficient resource utilization process the workload without compromising the QoS objectives under the constraint of less computing resources to be used. The aim of achieving the energy efficiency optimization not only to maximize the monetary benefits but it is worldwide pressure from the academia is to reduce the power consumption in the DCs for environmental safety. For achieving best trade-off between energy efficiency and QoS various VM consolidation techniques have been proposed [33][34][35]. Many of the researchers proposed resource allocation techniques that increase the resource utilization. The resource utilization helps the cloud providers maximize the revenue by assisting several customers as possible and reducing the operational expenditure by consolidating the workload on fewer machines. The relation of energy consumption and resource utilization is directly related and the QoS depends on the response time, latency, and throughput in cloud computing [36].

C. Allocation Method

According to the resource allocation method used, the resource allocation strategies can be categorized into two different categories such as Power-aware and Thermal-aware. The objective of power-aware strategies is to maximize energy efficiency with maintaining the rate of SLA violation [37]. The power-aware strategies can be applied in hardware and software and are also helpful to reduce energy expenditure. Resource utilization and workload history information is used to make the power-aware decision to reduce the energy wastage. While on other side thermal impact due to the workload placement is forecasted in thermal aware resource allocation techniques to keep the operating system into the operating temperature [38].

The thermal-aware resource allocation technique takes the temperature as the key consideration for allocation the cloud DC resources. The objective of thermal aware strategies is to reduce the inlet temperature that minimizes the significant cost of cooling.

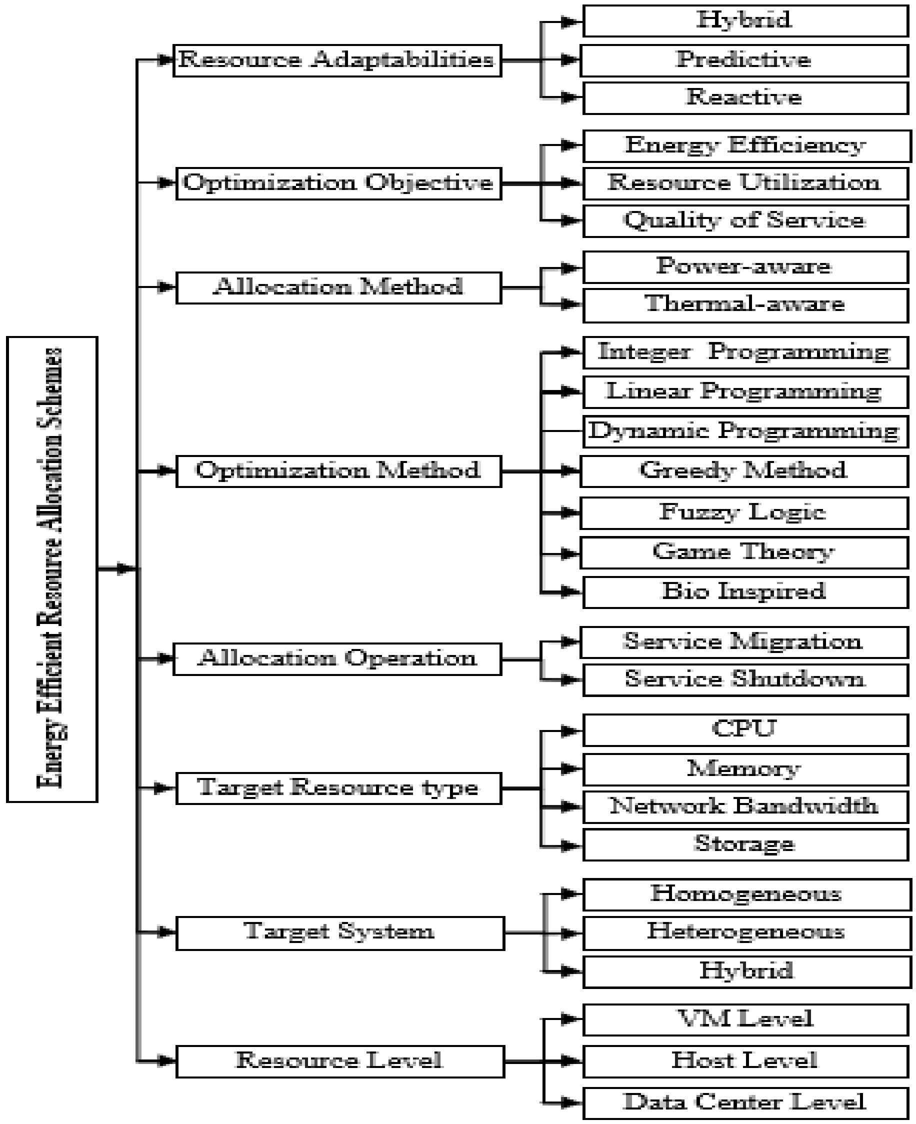

Fig.3. Taxonomy of Energy Efficient resource Allocation Schemes

Thermal- aware resource allocation and scheduling strategies keep all processing elements and data storage components within the temperature threshold limit to increase the energy efficiency and maintain the SLA by ensuring the availability of the system [39].

D. Optimization Method

Various optimization and mathematical models have been used by different researcher for resource allocation techniques. The optimization method metric represents the mathematical model practiced in various resource allocation strategies to find out the optimal or near-optimal solution for maintaining best balance between energy and SLA [40]. The problem of resource allocation cannot be solved by using the polynomials time algorithms, due the unpredictability nature, or fluctuating workload of emulative cloud user. The various optimization models have been found in energy efficient resource allocation schemes such as integer programming [41][42], liner programming[43], statistical optimization models [8][44][45], greedy optimization[46][47], greedy, fuzzy and bio-inspired optimization[48][49][50][51]. However, statistical optimization and bio-inspired models are costly in terms of mathematical computation in comparison of greedy, liner, integer optimization models; this is due to the number of iteration and computational complexity and stochastic process used to calculate the final result.

E. Allocation Operation

Service migration is the process of relocating the local data of an executing task instance from one physical machine to another physical, but the destination server must have sufficient amount of resources. The service migration features of virtualization enables the distribution of workload dynamically from over loaded server to under loaded server, [52]. Various energy-aware strategies have been developed to spot out overload and under loaded servers in order to maximize the energy efficiency. The aim of service shutdown is to free up active servers to reduce the power consumption in DCs but before switching off the servers, all the executing applications must be consolidated and migrated to others servers [53][54][55]. However, aggressive service migration, service shutdown and consolidation lead to performance degradation.

F. Target Resource Type

The target resource type metric represents the granularity level and types of resources allocated to user request. In CDCs heterogeneous types of resources are used to execute the user's workloads such as VMs, CPU, Memory, Storage, bandwidth and different networking switches. The target resource type field represents the pivot resources for which the resource allocation schemes are designed by researchers with the purpose to make the cloud DC green. At present current resource allocation schemes distribute the resources at various granularity levels. Some of the researcher have been developed the resource allocation strategies by considering CPU as the main component that consumes maximum amount of power of a physical machine in cloud DCs. The resources are allocated to the user application after optimizing the CPU utilization history of the PMs [35] [49] [58]. Some of the energy efficient resource management have been developed on the basis of multifactor i.e. CPU and Memory and improves the energy efficiency significantly [48] [61]. To allocate the resources in more energy efficient way some of the schemes have been developed by accounting CPU and network bandwidth, so that the VM allocation can be reallocated under the expected energy and SLA violation rate [62]. In the same way VM allocation shows the VM's placement on a PM. Likewise, storage allocation signifies the allocation of storage resources to VMs.

G. Target System

The target system attribute indicates the categorization of underlying servers for which the resource allocation schemes have been constructed. The servers in DCs are categorized into three different categories homogeneous, heterogeneous and hybrid. Homogeneous clusters consist of servers with identical capacities and performance level, while the heterogeneous clusters consist of servers with different resource capacity and performance grades. The hybrid cluster is the collection of homogeneous and heterogeneous ones [55]. Hadi khan et al. proposed a green power management for the distributed consolidation of the VMs using game theory and conducted experiments on the hybrid infrastructure to show that game converges to a Nash equilibrium after a finite number of migration, so that convergence time is bounded. Mohan raj et al developed a thermal aware adaptive heuristics for energy efficient consolidation of virtual machine for achieving energy efficiency. The targeted IaaS cloud DC is hybrid [56].

H. Resource level

The resource level attribute indicates kind of stages on which the resource allocation has been developed in different approaches. According to the literature survey the resource allocation generally found on three levels namely VM level, Host Level and DC level. Cho et al allocated the used task to the VM for making the proper load balancing and power optimization and proposed and proposed a dynamic resource allocation mechanism for energy optimization [64]. To allocate the available resources to the cloud workload in power efficient way many of the VM placement or allocation policies such as FFD, MBFD, PABFD, GPABFD, OBFD and SABFD have been found [56][57][58][59][60]. These VMs are directly mapped on to the hosts level (must have sufficient amount of resources) by using the different optimizations approaches like greedy [46], game theory [56], integer programming [41][42], liner programming[43]. In CDCs compute resources are the main key point where the energy is consumed most of the time. However, it is not always true as the networking switches, server racks and cooling systems also employ the significant amount of power [28][5]. In order to minimize the overall energy consumption at DCs the researcher allocated the resources by considering the networking switches, server racks, and cooling components of the DCs. Sina proposed a structure-aware VM consolidation strategy to improve the energy consumption of both the compute and cooling system. The proposed system enhance the energy efficiency of the overall DC by developing algorithms that takes into account DC configuration [63].

IV. CONCLUSION

The worldwide establishment of huge CDCs has led to the enormous amount of energy consumption and increase in CO_2 emission. The present paper briefly investigated the relationship between energy and power and explored the concept of resource allocation with energy efficiency and QoS. The paper discussed the significance of resource allocation in achieving energy efficiency and reviewed various existing resource allocation strategies. Finally, a taxonomy of energy efficient resource allocation schemes based on various attributes is presented. Energy efficient allocation of cloud DC resources to the user application is a challenging task and paves the way of further investigations for the researchers. Further, new methods for energy efficient resource allocation can be proposed by incorporating the techniques based on machine learning and soft computing.

REFERENCES

[1] Yue Gao, "An Energy and Deadline a Aware Resource Provisioning, Scheduling and Optimization Framework for Cloud Systems", International conference on Hardware/Software Codesign and system synthesis (CODES+ISSS), Montreal, QC, Canada 2013

[2] Arm burst M, Fox, "A view of Cloud computing". Vol. 53, Issue.4, pp:50-58.

[3] Zoltan Adam "Rigorous results on the effectiveness of some heuristics for the consolidation of virtual machines on cloud data center" Vol **51**, pages 1-6, 2015

[4] S Ali "Profit-aware DVFS enabled resource management of IaaS cloud". Vol. 10, Issue 2, pp-237-247, 2013

[5] Nasrin Ak., "Energy aware resource allocation of cloud data center: review and open issues" Vol. 19. Issue 3, pp-1163-1182, 2016

[6] Abdul Hameed "A Survey and taxonomy on energy efficient resource allocation techniques for cloud computing systems". Vol. 98. Issue 7, pp-751-774, 2016

[7] Anton Beloglazov: A taxonomy and survey of energy efficientDCs and cloud computing systems." Adv comput, Vol.82, pp :47-111, 2011

[8] Anton Beloglazov, Rajkumar,"Energy Efficient Resource Management in Virtualized Cloud Data Centers" International Conference on Cluster, Cloud and Grid Computing, Melbourne, VIC, Australia 2010

[9] S. Singh, I chana, "QoS-aware automatic cloud computing for ICT". International conference on information and communication technology for suitable development, pp: 569-577, 2015.

[10] Marston S, "A Cloud Computing. The business perspective", Decision support system, Vol 51, Issue 1, pp -176-189.

[11] Tarandeep Kaur " Energy Efficiency techniques in cloud computing" ACM Computing Surveys, Vol.48, No.2, Article 22, 2015

[12] Renewable Energy Outlook. 2013, "World Energy Outlook published by International Energy Associa- tion". Retrieved on August 5, 2014

[13] Tarandeep Kaur and Inderveer Chana, "Energy Efficiency Techniques in Cloud Computing: A Survey and Taxonomy", Vol. 48, Issue 2, pp 46, 2015.

[14] Koomey J, "Estimating total power consumption by servers in the US and the world". Lawrence Berkeley National Laboratory, Analytics Press, 2007

[15] Singh T, "Smart metering the clouds". In: 18th IEEE international workshops on enabling technologies: infrastructures for collaborative enterprises, pp 66–71

[16] J Baliga, "Energy consumption in optical IP networks". J Lightweight Technol Vol.27, Issue 13, pp:2391–2403, 2009

[17] O Tamm : "A Eco-sustainable system and network architectures for future transport networks". Bell Labs Tech J, Vol. 14, Issue 4, pp:311-327, 2010

[18] A Vukovic "DCs network power density challenges". J ASHRAE Vol. 47, pp:55–59, 2005

[19] J Liu " Challenges towards elastic power management in internet datacenters". International conference on distributed systems, pp: 65–72

[20] JA Paradiso, "Energy scavenging for mobile and wireless electronics". Pervasive Compute Vol.4, Issue, pp : 18-27

[21] Cook G, Horn J "How dirty is your data". GreenPeace International, Amsterdam, 2011

[22] Satveer , Mahendra Singh Aswal "A Comparative Study of Resource Allocation Strategies for a Green Cloud", International Conference on Next Generation Computing Technologies (NGCT-2016) Dehradun, India, 14-16 October 2016

[23] S Singh, I Chana, "Cloud resource provisioning: survey, status and future research directions" Knowl Inf Syst, Vol. 59, pp: 1005-1069

[24] Yousafzail,"Cloud Resource Allocation Schemes:Review, Taxonomy, and Opportunities", Knowl Inf Syst, Vol. 50, Issue-2 pp :347–381, 2017

[25] V.P Anuradha "Surey on Resource Allocation Strategies in Cloud Computing",InternationalConference on Information Communication and Embedded Systems (ICICES2014), Chennai, India ,IEEE.

[26] Morshedlou "Decreasing impact of SLA violations: a proactive resource alloca- tion approach for cloud computing environments". IEEE Trans Cloud Comput Vol. 2, Issue 2, pp:156–167, 2014

[27] Hussin,"Efficient energy management using adaptive reinforcement learning-based scheduling in large-scale distributed systems", International Conference on Parallel Processing In: ICPP, Taipei City, Taiwan, pp 385–393, 2011

[28] Lee, "Resource Allocation and Scheduling in Heterogeneous Clooud Enviroments" Ph.D. dissertation, Univ. California, Berkeley, Technical Report No UCB/EECS-2012-78, spring 2012

[29] TVT. Duy, "Performance evaluation of a green scheduling algorithm for energy savings in cloud computing". International Symposium on Parallel and distributed processing, workshops and PhD forum (IPDPSW), Atlanta, GA, USA Atlanta, GA, USA pp 1–8, 19–23, 2010

[30] Mezmaz, "A parallel island-based hybrid genetic algorithm for precedence-constrained applications to minimize energy consumption and makespan". International Conference on Grid Computing, Brussels, Belgium, pp 274–281, 2011

[31] Y. Chen, "Minimizing data center SLA violations and power consumption via hybrid resource provisioning".Second international green computing conference (IGCC), pp 1–8, 2011

[32] E. Kalyvianaki,"Resource provisioning for virtualized server applications". Technical Report UCAM-CL-TR-762, Computer Laboratory, University of Cambridge, 2009

[33] Anton Beloglazov "Adaptive Threshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Data Centers": MGC , Bangkore, India ISBN: 978-1-4503-0453-5 ,

[34] Mohan Raj: "Heterogeneity and thermal aware adaptive heuristics for energy efficient consolidation of virtual machines in infrastructure clouds" Journal of Computer and System Sciences, Vol.82, Issue. 2, pp: 191-212

[35] Fahimeh, "Energy Aware Consolidation Algorithm based on K-nearest Neighbor Regression for Cloud Data Centers", International Conference on Utility and Cloud Computing, Dresden, Germany 2013

[36] Nasrin Akhter, "Energy aware resource allocation of cloud data center: Review and open issues" Cluster Compute, New York, Vol 19, Issue, 3,pp: 1163-1182 2016

[37] More,"Energy-efficiency in cloud computing environments: towards energy savings without performance degradation". International Journal of computer Applications, Vol. 1, Issue. 1, pp:17–33, 2011

[38] C, Reid, "Coordination of energy efficiency and demand response". Environmental Energy Technologies Division, Berkeley National Laboratory, LBNL-3044E, 2010

[39] N, Scherer, "Thermal-aware workload scheduling for energy efficient data centers". International conference on autonomic computing (ICAC) Washington, DC, USA, pp 169–174.

[40] Heger, "Optimized Resource Allocation & Task Scheduling Challenges in Cloud Computing Environments". 2010, dheger@ dhtusa. com

[41] I, Aida, "Applying double-sided combinational auctions to resource allocation in cloud computing". International symposium

on applications and the internet. Seoul, South Korea, pp 7–14, 2012
[42] MM. Mashayekhy "Truthful greedy mechanisms for dynamic virtual machine provisioning and allocation in clouds". IEEE Transactions on Parallel and Distributed Systems Vol: 26 , Issue: 2 , pp: 594–603, 2015
[43] Y, Niyato, "An auction mechanism for resource allocation in mobile cloud computing systems". International Conference on Wireless Algorithms, Systems, and Applications, pp76-87, 2013
[44] W-Y. Lin, "Dynamic auction mechanism for cloud resource allocation". International conference on, cluster, cloud and grid computing (CCGrid),IEEE, Melbourne, VIC, Australia, pp 591–592, 2010
[45] Li Z, "An anti-cheating bidding approach for resource allocation in cloud computing environments". Journal of Computational Information Systems, Vol 8, Issue: 4, pp:1641–1654, 2012
[46] Dharmesh Kakadia: "Network-aware Virtual Machine Consolidation for LargeDataCenters" NDM, proceding of the third international workshop on Network aware data management, 2013
[47] S. Grosu, "A combinatorial auction-based mechanism for dynamic VM provisioning and allocation in clouds". IEEE Trans Cloud Comput, Vol 1, Issue 2, pp:129–141, 2013
[48] Ajit Singh, "Cluster Based Bee Algorithm for virtual Machine Placement in CloudDC" Journal of Theoretical and Applied Information Technology, Vol. 57, 2013.
[49] K.Mukkherjee,"Green Cloud:An Algorithmic Approach" International Journal of Computer Applications" (0975-8887) Vol. 9,2010
[50] J. Wang "An auction and league championship algorithm based resource allocation mechanism for distributed cloud". In: Wu C, Cohen A (eds) Advanced parallel processing technologies, Vol. 8299., pp 334–346, 2013
[51] C, Wang, "A cloud resource allocation mechanism based on mean-variance optimization and double multi-attribution auction". In: Hsu C-H, Li X, Shi X, Zheng R (eds) Network and parallel computing, vol 8147, pp 106–117, 2013
[52] W-L, Xie, "Thermal-aware task allocation and scheduling for embedded systems". Proceedings of the conference on design, automation and test in Europe, vol 2, pp 898–899, 2005
[53] S, Kansal, "Energy aware consolidation for cloud computing". In: Confer- ence on power aware computer and systems, San Diego, California ,2008
[54] A, Ahuja P, Neogi A (2008) "pMapper: power and migration cost aware application placement in virtualized systems". International conference on middleware, pp 243–264, 2008
[55] R, Schwan K,"VirtualPower: coordinated power management in virtualized enterprise systems". In: 21st ACM SIGOPS symposium on operating systems principles, Vol.41, Issue. 6, pp: 265–278, 2007.
[56] Hadi Khani, "Distributed consolidation of VMs for power efficiency in heterogeneous cloud data centers", Journal of computers and Electrical Engineering, Vol. 47 Issue C, pp: 173-185, 2015
[57] Anton Beloglazov, "Optimal Online Deterministic Algorithms and Adaptive Heuristics for Energy and Performance Efficient Dynamic Consolidation of Virtual Machines in Cloud Data Centers" Journal of concurrency and computation, Vol. 24, Issue. 13, 2012
[58] Anton "Energy Aware resource allocation heuristics for efficient management of cloud DCs" Future generation computing system VOl. 28, pp: 755-768, 2012
[59] Alfredo Goldman "Consolidation of VMs to improve Energy Efficiency in cloud Environments" 2015 XXXIII Brazilian Symposium on Computer Networks and Distributed Systems, Vitoria,Brazil, 2015
[60] Hui Wang "Energy-aware Dynamic Virtual Machine Consolidation for CloudDCs" IEEE 7th International Conference on Cloud Computing, Anchorage, AK, USA, 2014
[61] Bruno Cesar Ribas: "On Modelling Virtual Machine Consolidation to Pseudo-Boolean Constraints" J. Pavo´n et al. (Eds.) pp. 361–370, 2012.
[62] Dabiah Ahmed, "Energy-aware Virtual Machine Consolidation for Cloud Data" International Conference on Utility and Cloud Computing Centers" London UK , 2014
[63] Sina Esfandiarpoor "Sturecture-aware online VM Consolidation for datacenter energy improvement in cloud cloud computing" Computers and Electrical Engineering, Vol. 42, pp:74-75, 2015
[64] ChaoTung, "Green Power Management with Dynamic Resource Allocation for Cloud Virtual Machines", International Conference on High Performance Computing and Communications Banff, AB, Canada 2011

Authors Profile

Mr. Satveer pursed Bachelor of Science and and Master of Computer Application from Gurukula Kangri Vishwavidyalaya Haridwar, India in 2011 and 2014 respectively. He is currently pursuing Ph.D in Department of Computer Science, Gurukula Kangri Vishwavidyalaya Haridwar, India. His main research work focuses on Cloud Computing, Green Cloud Comuting, IOT.

Dr Mahendra Singh is currently working in the Computer Science department of Gurukul Kangri Vishwavidyalaya, Haridwar, India as Asst Professor. He has diverse and vast experience of more than 20 years in the field of teaching/system administration. He has got master degrees in Physics and Computer Science. He obtained his Ph.D. in Computer Science in 2007 in the field of interconnection networks. He has got published more than thirty research papers in different National/ International Journals and conference proceedings. He has authored three books. His areas of interest are parallel and distributed computing, AI, Networks and Green Computing. He is a life member of Computer Society of India(CSI) and Indian Science Congress.

