Cloud computing: a walk through in models, challenges, and energy solutions

N Gupta¹², K Gupta¹³, A Singh¹⁴, A Kaur¹⁵ and D Gupta¹⁶
¹Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India

Email: ²neha.gupta@chitkara.edu.in, ³kamali.singla@chitkara.edu.in,
⁴amanpreet.singh@chitkara.edu.in, ⁵amandeep.bhullar@chitkara.edu.in,
⁶deepali.gupta@chitkara.edu.in

Abstract. The paradigm of cloud computing in recent times has engrossed a lot of working populace in its service stack offerings. This considerable rise of human traffic landing onto cloud platform for service extraction has consequently led to exponential rise in energy consumption during dynamic and heterogeneous workload situations encompassing huge cloud infrastructure in form of CPU, Hard Disks, network devices etc. The energy released not only poses costs on its stakeholders but is affecting the environment due to rise in carbon footprints. As per available stats [1], the total power consumption of cloud has significantly arisen from 632 billion kilowatt hours in year 2007 to 163 billion kilowatt hours in 2020 which is adequate to leave a devastating effect on environment. Though research is being carried out in this direction, still efficacious measures are far away from its pragmatic implementations. Therefore, the research study conducted here discusses about existing solutions in the direction of energy management in cloud computing. The latter half of paper presents abrief discussion on gaps and future directions.

Keywords: Cloud Computing, Energy Management, Virtualization, Bin-packing, VM migration.

1. Introduction to Cloud Computing

Since ages, the services extended by the model of Cloud Computing are extensively being utilized by the user entities for advancing into technological era for reaping its benefits in day-to-day activities. The model is a pay-per-use utility model that comprises of a collection of datacenters for storage, servers, networking services and other resources enabling users to use them on internet.
Existence of cloud is from 1950 with changes from static implementation to dynamic implementation on internet [1]. Rather than purchasing and maintaining all computing resources at their own place, companies prefer to rent access to all resources from simple application like Gmail to complicated natural language processing & Artificial intelligence from a cloud service provider [2]. Its architecture refers to how users connect to the physical computing resources provided by cloud service provider. Cloud computing comprises of two parts- frontend and backend platforms [4]. Deployment models of cloud helps user to identify cloud according to his requirements. This model divides cloud based on usage, storage size, accessibility, and proprietorship. According to Deployment model, cloud can be of four types: private, community, public and hybrid cloud [5]. Service Models tells the services which can be provided by cloud. The term XaaS (Everything as a service) can be used for services provided by cloud vendor [8]. Depending upon the services, cloud can be categorized as SaaS, PaaS, and IaaS [6] [7] as depicted in Figure 1.

Cloud also faces some challenges which needs to be resolved, which primarily includes, Data Management and Resource Allocation for managing unlimited customer’s access for resource allocation and utilization, ensuring security and privacy, balancing of load to manage under and over utilization of physical and virtual machines and energy consumption as the greatest threat for environment and cost as shown in Figure 2. Energy requirement of data centers depends on its physical layout and can range from few kilowatts of energy up to tens of megawatts or more. Energy efficiency of data center can be measured as power usage effectiveness (PUE), which is the ratio of total input power divided by total power used by physical resources [3]. A walk through in the challenges enlisted above has derived an

![Figure 1: History, Deployment Model and Service Stack of Cloud [27][28]](image-url)
observation that the energy consumed in cloud paradigm is not only posing escalated cost challenges to the service providers, but is also causing health hazards to the environment which is a major concern in today’s technological era. Thus, the next section briefly embarks on energy management in cloud environment.

Figure 2: Energy Representation in Cloud [29]

2. Energy management in cloud computing

Data storage and analysis has become major task of company’s total workload. To enhance the efficiency of storage and analysis, all companies are shifting their data on cloud where they get the entire infrastructure like servers, storage, CPU computing power etc. on pay-per-use basis instead of maintaining infrastructure at their own place. Cloud data centers need to be functional all time to maintain data of their customers, so they are consuming large amount of energy 24/7. This energy requirement is so high that it becomes major challenge for cloud service providers as it is not only increasing the cost of maintaining data centers but also affecting environment by continuously emitting carbon footprints.

According to the statistics, electricity consumption in data centers is 2 % of the total energy consumption globally and it can rise to 8% by 2030 [11]. Carbon emission by data centers is about 0.3% of total carbon emission. The concept of energy management can be realized by considering the role of parameters in the domain of cloud that can contribute to strategies pertaining to reduction in energy consumption. The existing energy management solutions uncovered two important aspects. First embarks upon Bin-Packing where-in the idle server machines are turned off by shifting their load onto other machines. But another puts emphasis on dealing with overheads incurred at the time of migration. Mitigating these migration overheads can significantly contribute to reduction of energy. An important aspect that greatly contributes energy management is virtualization, and thus the aspects have been elucidated in subsequent section.

3. Bin packing: virtualization and system model

Number of physical machines in data centers are limited so cloud service provider uses the technique of virtualization through which m number of virtual machines can be created on each physical machine. By using virtualization, the total number of users who can use cloud services are m*n [12].Sometime, one physical machine gets overloaded, so shifting of VM is required from one physical machine to other, this is called placement of VM. For the mathematical representation of placement, lets S is the number of Servers, V is the number of virtual machines,J is the number of jobs, S*V be the placement matrix denoted
by $F_{old}$ and $F_{new}$ and the matrix values for old and new placement, $f_{xy}$ is equal to 1 if $v_x$ is hosted on server $s_y$, otherwise 0. $Req\_CPU_k(t)$ is the computing capacity on server which is required by VM $v_k$. $CPU_x$ is the total energy capacity by server $s_x$ and $P_{max}^x$ is the maximum power consumed by server $s_x$.

$$U_x(F,t) = \sum_{k=1}^V f_{yk} \times (Req\_CPU_k(t)/CPU_{yk})$$ (1)

$$\sum_{k=1}^V f_{yk} \times (Req\_CPU_k(t)/CPU_{yk}) \leq 1$$ (2)

$$P_x(F,t) = 0.7P_{max}^x + 0.3P_{max}^x \times U_x(F,t)$$ (3)

$$Energy(F,t_1,t_2) = \sum_{x=1}^S \int_{t_1}^{t_2} P_x((F,t))dt$$ (4)

$$Comm(F, t_1,t_2) = \sum_{m=1}^V \sum_{k=1}^V \sum_{x=1}^S \sum_{y=1}^S c_{mk}^{xy} f_{mk} f_{ky} d_{xy}$$ (5)

$$Impl(F_{old}, F_{new}) = \sum_{k=1}^V \sum_{x=1}^S \sum_{y=1}^S (V_k) f_{kx}^{old} f_{ky}^{new} d_{xy}$$ (6)

Equation (1) tells the utilization of server as a function of placement F and time t. Equation (2) tells that utilization of server should not be greater than 1. Equation (3) tells the power requirement of server $s_x$. Equation (4) tells total energy required by all servers for placement F and duration $t_1$ and $t_2$. Equation (5) tells total network load by all servers for placement F and duration $t_1$ and $t_2$. Equation 6 tells implementation cost of moving from $F_{old}$ to $F_{new}$.

VM migration plays important role in Cloud computing to decrease overall energy consumption in cloud data centers. Various algorithms have been developed in this topic which are discussed in next section.

4. Virtual machine migration

Virtual Machine migration is the procedure of shifting VM from one server (overloaded or maximum energy consuming) to other server (under loaded or least energy consuming) to decrease the overall energy consumption of data centers by turning off the idle machines. The technique is commonly known as Bin-Packing and inculcates the following heuristics for its pragmatic implementation:
Power and Computing capacity- Aware Best Fit Decreasing (PCA-BFD)
This algorithm works in same manner as previous algorithm BFD but servers are arranged in ascending order of $P_{\text{max}}/\text{CPU}_x$. Maximum computing capacity required VM is assigned on Server with minimum $P_{\text{max}}/\text{CPU}_x$ ratio. Remaining working of algorithm is exactly same as in BFD. **Algorithm:**

1) For $x=1$ to $S$
2) Compute $\text{CPU}_x, P_{\text{max}}$ for all servers
3) Create VM list with computing capacity in descending order
4) Create Server list with $P_{\text{max}}/\text{CPU}_x$ in ascending order
5) For $i=1$ to $V$
   a) If $\text{T}_{\text{CPU}}i<=\text{CPU}_i$ then $s_i=\text{v}_i$
6) Switch off the servers on which no VM is assigned and delete the server from server list.
7) Calculate Energy consumption of complete system

EUBFD-Energy-Utilized BFD Packing Algorithm
In this algorithm, VM list is arranged in ascending order of computing capacity and Server list is arranged in ascending order of Power consumption. If two servers have same power consumption, they are arranged according to computing capacity. After creating VM and server list, assignment of VM on server starts and continue for all VM in the list. **Algorithm:**

1) For $x=1$ to $S$
2) Compute $\text{CPU}_x, P_{\text{max}}$ for all servers
3) Create VM list with computing capacity in descending order
4) Create Server list with $P_{\text{max}}/\text{CPU}_x$ in ascending order
5) For $i=1$ to $V$
   a) If $\text{T}_{\text{CPU}}i<=\text{CPU}_i$ then $s_i=\text{v}_i$
6) Switch off the servers on which no VM is assigned and delete the server from server list.
7) Calculate Energy consumption of complete system

Low Perturbation Bin Packing Algorithm (LPBP)
This algorithm focuses on migrating VM from maximum energy consumed server to least energy consumed server. The energy consumption of each server is calculated and then these servers are arranged in decreasing order of energy consumption. $T_{\text{CPU}}x$ is the sum of energy required by all VMs on server $x$.

**Algorithm:**

1) For $x=1$ to $S$
2) Compute $\text{CPU}_x, P_{\text{max}}$ for all servers
3) Random $\{s_x\}$
4) For $x=1$ to $S$
   a) Calculate $U_x(F,t)$ and $P_x(F,t)$
5) Arrange servers in decresing order of $P_x(F,t)$.
6) Let $T_{\text{CPU}}x$ (total energy of max. energy consumed server)
   a) $T_{\text{CPU}}x=\sum_{i=1}^{\text{K}}\text{Req}_{\text{CPU}}i$/CPU$_x$ for k number of VMs on server $s_x$
   b) If $\text{T}_{\text{CPU}}y<\text{CPU}_x$ then $s_y<s_x$
7) Calculate Energy consumption of complete system

Best Fit Decreasing (BFD)
This algorithm checks the computing capacity of VM before assigning it to the server. VM list is arranged in descending order of their computing capacity and Server list is arranged in ascending order of their computing capacity. Top VM from VM list means maximum computing capacity required VM is assigned to top server from server list means least computing capacity providing server. This process follows till all VMs get assigned to server. **Algorithm:**

1) For $x=1$ to $S$
2) Compute $\text{CPU}_x, P_{\text{max}}$ for all servers
3) Create VM list with computing capacity in descending order
4) Create Server list with computing capacity in ascending order
5) For $i=1$ to $V$
   a) If $\text{T}_{\text{CPU}}i<=\text{CPU}_i$ then $s_i=\text{v}_i$
6) Switch off the servers on which no VM is assigned and delete the server from server list.
7) For $x=1$ to $S$
   a) Calculate $U_x(F,t)$ and $P_x(F,t)$
   b) Calculate Energy consumption of complete system
The comparative analysis of the existing Bin-Packing techniques has been conducted in C and the results are presented in Table 1 and Figure 3 presented below. The simulation results reveal the Energy-Utilized BFD can considerably reduce energy in comparison to other existing Bin-Packing techniques.

**Table 1. Evaluation of Bin Packing Techniques**

| No. of Simulated Servers | No. of Virtual Machines | Energy Spent in kWh |
|--------------------------|-------------------------|---------------------|
|                          |                         | LPBP    | BFD    | PCA-BFD | EUBFD |
| 5                        | 8                       | 120.78  | 302.73 | 110.16  | 104.78 |
| 15                       | 38                      | 1383.53 | 2378.96| 1068.23 | 960.59 |
| 25                       | 60                      | 4445.56 | 5912.83| 5326.81 | 3345.23|
| 35                       | 60                      | 7621.5  | 9605.72| 6578.24 | 5747.33|
| 45                       | 85                      | 5976.9  | 8784.8 | 4758.1  | 4660.9 |

**Figure 3: Comparative study of Bin Packing Techniques**

5. Related work

As energy consumption in cloud data centers has become major threat over customer investments and environment, continuous work is going on to find the effective solution. Table 2 presents the overview of similar research study conducted in this direction:

**Table2. Overview of research to optimize the energy**

| Algorithm/Policy proposed                              | Performance metrics                          | Simulation Tool | Year |
|--------------------------------------------------------|----------------------------------------------|-----------------|------|
| Enhanced Bin-packing algorithm                         | VM Migration, Bin Packing                    | C               | 2016 |
| Algorithm “VM scheduler”                               | Resource Utilization                         | CloudSim        | 2016 |
| Minimum Average Utilization Difference policy          | Virtual machine consolidation                | CloudSim        | 2016 |
| Algorithm “Power aware best fit (PABF)”               | Virtual machine consolidation                | CloudSim        | 2017 |
As shown in above table, VM consolidation is very effective technique for reducing energy consumption. Energy can be saved maximum 50% by using VM consolidation technique [25]. VM consolidation helps in using some servers 100% while shutting down of idle servers. Idle server consumes about 70% of the energy which it consumes when fully utilized [25]. Live VM migration is the process to migrate VM from one physical machine to other. Each VM has some properties which needs to be transferred on destination machine like memory pages, virtual disks, and network connectivity. Memory data migration deals with migration of memory pages from source PM to destination PM. Storage Data Migration deals with migration of blocks of storage which occurs inter-data centers. This is different from memory migration as size of data to be transferred is large. Network Connection Continuity is preservation of network connection between VM and user. In Layer 2, same network can be used by extending LAN to multiple data centers. Layer 3 provides some methods like IP tunneling [26] to forward the packets from source to destination through tunnel.

6. Research gaps and future discussion

Although, Live VM migration is considerably reducing the energy consumption still extra overheads are being incurred that are contributing to increase in energy, hence needs to be addressed. For each live VM migration, VM is transferred from source host to destination host along with memory state, virtual disk state and network connectivity with user. Two most important parameters for performance of migration are:

- **Downtime**- During migration, the host machines involved in the process of migration gets halted and thus the downtime needs to be reduced as during migration they are contributing to energy consumption but effectively not performing any computing or related processes.
- **Total migration time**- Total migration time needs to be reduced as it is consuming bandwidth; thereby, utilizing the resources and leading to rise in energy consumption in data centers. It depends on the total amount of memory transferred from source to destination server and allocated bandwidth or link speed.
Inclination should be driven towards minimizing total amount of data transferred during migration by reducing identical pages and unused pages, by sending most modifiable pages (dirty pages) at the end etc. Less research work has been carried out in this direction. More robust mechanism needs to be proposed for handling atomicity and convergence in live VM migration. Although, Bin-packing solutions are existing, still more robust heuristic needs to be generated for reducing energy consumption considerably.

7. Conclusion

Although cloud is supplementing technological solutions assisting in day-to-day endeavors, still the energy release in data centers is a major issue that needs to be addressed. Comprehending solutions for Bin-Packing and live migration overheads can greatly reduce energy emissions in the environment. Virtualization is a boom that has contributed much for increasing resource utilization and extraction of services to a higher extent with the available resources. This research study has majorly focused on exploring the existing solutions in the direction of Bin-Packing and Live Migration. The simulation results show that EUBFD among the existing Bin Packing techniques is reducing energy consumption. For researchers inclined towards conducting research in this direction, the gaps are identified in research gaps and future discussion section.

8. References

[1] https://www.tutorialspoint.com/cloud_computing/cloud_computing_overview.htm.
[2] https://www.zdnet.com/article/what-is-cloud-computing-everything-you-need-to-know-from-public-and-private-cloud-to-software-as-a/.
[3] https://searchdatacenter.techtarget.com/definition/data-center.
[4] https://www.tutorialspoint.com/cloud_computing/cloud_computing_architecture.
[5] Savu L, (2011). Cloud computing: Deployment models, delivery models, risks and research challenges. 2011 International Conference on Computer and Management, CAMAN 2011.
[6] Hari Krishna B, Kiran S, Murali G, and Pradeep Kumar Reddy R, (2016). Security Issues in Service. Model of Cloud Computing Environment. Procedia Computer Science, 87, 246–51.
[7] https://www.tutorialride.com/cloud-computing/service-models-in-cloud-computing.htm.
[8] https://www.sam-solutions.com/blog/everything-as-a-service-xxas-definition-and-examples/.
[9] https://www.newgenapps.com/blog/top-10-cloud-computing-examples-and-uses.
[10] FatemiMoghaddam F, Ahmadi M, Sarvari S, Esiami MandGolkar A, (2015). Cloud computing challenges and opportunities: A survey. 2015 International Conference on Telematics and Future Generation Networks, TAFGEN 2015, 34–8. https://doi.org/10.1109/TAFGEN.2015.7289571.
[11] https://fortune.com/2019/09/18/internet-cloud-server-data-center-energy-consumption-renewable-coal/.
[12] Tziritas N, Xu C Z, Loukopoulos T, Ullah Khan Sand Yu Z, (2013). Application-Aware workload consolidation to minimize both energy consumption and network load in cloud environments. Proceedings of the International Conference on Parallel Processing, 449–57.
[13] https://shodhganga.inflibnet.ac.in/bitstream/10603/219788/14/14_chapter5.pdf.
[14] Gupta Kand Katiyar V, (2016). Energy Aware Virtual Machine Migration Techniques for Cloud Environment. International Journal of Computer Applications (0975 – 8887) 141(2).
[15] Alismail S M and Kurdi H A, (2016). Green Algorithm to Reduce the Energy Consumption in Cloud Computing Data Centres. SAI Computing Conference 2016(557–61).
[16] Deng D, He K and Chen Y, (2016). Dynamic virtual machine consolidation for improving energy efficiency in cloud data centers, *Proceedings of CCIS2016*.

[17] Guo L, Zhang Y and Zhao S (2017). Heuristic algorithms for energy and performance dynamic optimization in cloud computing. *Computing and Informatics, 36*, 2017, 1335–60, doi: 10.4149/cai_2017_6_1335.

[18] Han G, Que W, Jia G and Zhang W, (2017). Author’s Accepted Manuscript Resource-utilization-aware energy efficient server consolidation algorithm for green computing in IOT. *Journal of Network and Computer Applications*. https://doi.org/10.1016/j.jnca.2017.07.011.

[19] Gupta K and Katiyar V, (2017). Energy-Aware Scheduling Framework for resource allocation in a virtualized cloud data centre. *9*(2), 558–63.https://doi.org/10.21817/ijet/2017/v9i2/170902032.

[20] Sharkh M A and Shami A, (2017). An Evergreen Cloud: Optimizing Energy Efficiency in Heterogeneous Cloud Computing Architectures. *Vehicular Communications, February*.

[21] Diouani S and Medromi H, (2018). *Green Cloud Computing: Efficient Energy-Aware and Dynamic Resources Management in Data Centers*. July 10–14. https://doi.org/10.14569/IJACSA.2018.090717.

[22] Karuppasamy M and Balakannan S P, (2018). Energy Saving from Cloud Resources for a Sustainable Green Cloud Computing Environment. *7*, 95–108. https://doi.org/10.13052/jcsm2245-1439.718.

[23] Jeba J A, (2019). Towards Green Cloud Computing an Algorithmic Approach for Energy Minimization in Cloud Data Centers. *9*(1). https://doi.org/10.4018/IJACAC.2019010105.

[24] Panda S K and Jana P K, (2019). An energy-efficient task scheduling algorithm for heterogeneous cloud computing systems. *Cluster Computing, 22*(2), 509–27.

[25] Derdus K M, Omvenga VandPatrick O J, (2017) Energy Consumption in Cloud Computing Environments.

[26] Zhang F, Liu G, Fu X and Yahyapour R, (2018). A Survey on Virtual Machine Migration: Challenges, Techniques, and Open Issues. *IEEE Communications Surveys and Tutorials, 20*(2), 1206–43.

[27] Alex Mu-HsingKuo (2011). Opportunities and Challenges of Cloud Computing to improve Health Care Services. *JMIR publications, 13*, Jul-Sep.

[28] https://www.researchgate.net/figure/Cloud-Computing-Service-Model_fig1_259865832.

[29] Barbulescu M, Grigoriu RO, Neculoiu G, Haleu I, Sândulescu V C, Niculescu-Faida O, Marinescu M, and Marinescu V (2013). Energy efficiency in cloud computing and distributed systems. *Proceedings - RoEduNet IEEE International Conference*. https://doi.org/10.1109/RoEduNet.2013.6714197.