

Experimenting with Energy Efficient VM Migration in IaaS Cloud: Moving Towards Green Cloud

Riddhi Thakkar^(✉), Rinni Trivedi, and Madhuri Bhavsar

Institute of Technology, Nirma University, Ahmedabad 382481, Gujarat, India
{15mcen27,14mcei28,madhuri.bhavsar}@nirmauni.ac.in

Abstract. Increasing demand for Cloud infrastructure and services leads to the challenges for management and maintenance of large data Center. Data center is fully equipped with huge number of resources. Those resources consumes energy in spite of their partial or full utilization. As a result data center consumes lots of energy, which in turn increases the total cost of operation and carbon footprint in environment. These concern leads to “Green Computing”, i.e. to reduce total operational cost, Carbon Footprint in environment and efficient usage of the computing resources. In data center main processing element is virtual machine (VM), which is an instance of computing and storage resources, handles computational processes. Hence, it is important to reduce energy consumed by VM. As the workload distribution is varying in data Center as per the need, the number of VMs configured in the host are uneven, but host consumes maximum energy every time, irrespective of the workload. This leads to wastage of computational resources. This paper is intended to analyze such issues and specifically prove an algorithm which, significantly reduces energy consumption in data center, while ensuring SLA, when VM is in migration from one host to another in the data center.

Keywords: Energy efficiency · Green cloud · Vm migration

1 Introduction

Cloud computing is a computing model which provides access to shared pool of resources on demand, which can be swiftly provided and released with less management efforts and interaction of service provider and charged as much use [2]. It provides high availability and facilitating some important characteristics: Elasticity, on demand self service, Resource Pooling, and Network access. Cloud Computing Supports three service models IaaS, SaaS and PaaS, and uses virtualization technology to deliver scalability and elasticity to IT services. This facilitates to avoid over-provisioning and under-provisioning and to cater a new business utilities based on new ways of operating [3]. Cloud Computing is

widely used, but Resource Utilization, Security, Portability, and Interoperability are still major issues to deal with.

Computing resources inside data center consumes huge amount of energy, produces thousands of ton carbon dioxide and harms the atmosphere [5]. Energy is also required for functioning of the cooling system. Extra 0.5 W is required by the computing resources for each watt of power consumption for cooling [9]. According to [1] if any solution is not applied to reduce carbon emission by the data center, then it will be triple 2020 as compared to 2002. As indicated by the European Union, for keeping global temperature grow below 2°, it is necessary to drop the carbon emission up to 15% to 30% [11]. So, Consumption of Energy by Cloud Infrastructure as well as Carbon dioxide emission by that have been an important environment consideration. This will lead to use the Eco inviting computing called “Green Cloud” for minimizing the computing cost and to decrease the environmental destruction. Green means Environment-friendly and in other words utilization of resources.

Energy Efficiency can be achieved in Cloud Computing by providing some characteristics [9]: Improved Resource Utilization, Portability: VMs can be migrated to other physical host to reduce energy consumption, Elasticity, upgrading the running time of algorithms in application, DVFS and Virtualization of resources. In DVFS, the amount of voltage given to resource is decreased, when it is not completely utilized. The disadvantage is, it cannot apply to any resource of the cloud other than CPU, and frequency/voltage can be set to a limited number of states. Virtualization allows the user to create multiple VMs on single hardware and to run different OS on it. It provides heterogeneity, improves performance isolation and fault isolation by migrating VMs from one host to other host by Hot or Cold migration. In summary, it is necessary to utilize the cloud resources optimally, by migrating and consolidating VMs in less number hosts. This will decrease the carbon footprint and energy consumed by cloud resources.

2 Domain Analysis

Sherif et al. [4] investigated various techniques for migration and various stages in that techniques. VM migration parameters are classified into two parts: Static and Dynamic, based on their effect on VM migration. Static effect parameter has an unavoidable migration overhead. Dynamic effect parameter influences only the transfer of process, which are typically associated with the VM specification and its hosted application. The parameters that are strongly affecting the VM migration behavior are Page Dirty Rate and Link Speed. The author concluded that migration in XEN cannot escalate with a high-speed link.

Anton et al. [6] specified some key issues that must be undertaken for energy efficiency and performance. VM allocation policy is divided into two parts: The submission of a request for VM provisioning. For VM provisioning, changes are done in Best Fit Decreasing Algorithm. Optimization of VM allocation is performed in two steps: 1. To choose the VM for migration, 2. To choose an appropriate host for that VM. Author discussed RC, MMT, and highest potential

growth polices for choosing VM for migration. Paper shows, Minimum Migration policy perform less VM migrations and decrease SLA violations than other policies evaluated for simulation scenario.

Akshat et al. [13] have analyzed server workload of the large data center. To save energy, author experiment many attributes for the long and short term and conclude, there is an opportunity for saving power with help of off-peak metrics for categorizing application. While doing consolidation, if care is not taken then there will be possibilities of capacity violations. Two new consolidation approaches proposed: Peak Clustering based Placement (PCP) and Correlation Based Placement (CBP). These approaches use two metrics, first is an off-peak metric for categorizing and second is to guarantee that peaks don't cross limits. It is concluded, PCP provides better energy saving, decreased SLA violations and efficient load balance among active server.

Anton et al. [8] have investigated dynamic VM consolidation and single VM migration problem. In modern applications, because of variability in workload, VM placement is optimized online. For dynamic VM consolidation, author proposed some techniques, Median Absolute Deviation, Robust Local Regression, Interquartile Range, and Local Regression. These techniques use historical workload [7], generated from online services and web applications [4]. The goal of these techniques is to keep CPU utilization between upper and lower utilization threshold. For VM selection also, several policies discussed: MMT, RC, and Maximum Correlation (MC). From experiment, author shows that IRQ with MMT provides better energy reduction by reducing VM migration and SLA violations.

3 Energy Efficiency in IAAS Cloud

In data center, main processing element is VM, which is an instance of computing and storage resources, and does computational processes. So, VM plays important role in reducing energy consumption in data center. VMs are collected from an unutilized host. These VMs are consolidated into lesser number of host, and other hosts are switched off by migrating their load on active hosts. Hence, less number of host are active and energy is consumed by them only. In that way, overall energy consumption can be reduced in data center.

3.1 Proposed Architecture

Proposed architecture is intended for IAAS service. User submits a request for resource provisioning. Accordingly, SLA's are negotiated between Cloud Service Provider (CSP) and user. CSP does Service Scheduling and manages the account of the user. Monitoring kit is a heart of the system. It monitors power consumed by host and decides which host is over utilized or underutilized. It chooses VMs from over utilized host and consolidates them in less number of host by doing migration. VMs are configured on hosts. Each host is characterized by utilization of Million Instruction per Second (MIPS), Storage Capacity, network Bandwidth (BW) and RAM. User submits a request for provisioning

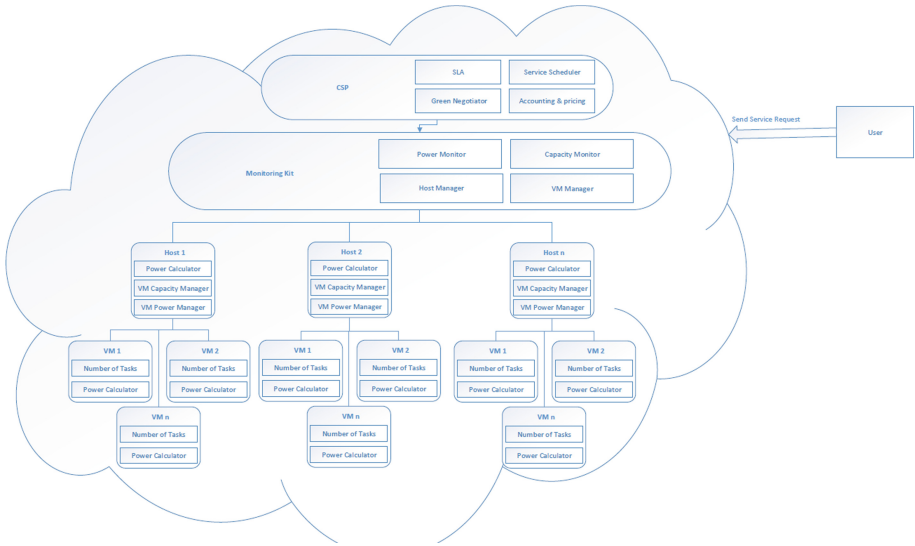


Fig. 1. Proposed architecture

of VMs, are allocated according to requested amount of MIPS, RAM, BW and Storage. While providing service to user CSP needs to take care of SLA.

3.2 Proposed System Model

Figure 2 shows, the snapshot of a system without implementation of energy efficient algorithm. In such system hosts are not fully utilized, there is less number of VMs are executing inside it then its capacity. When energy efficient algorithm is applied to the system, algorithm identifies underutilized host by comparing utilization parameter and collect the VMs from host. These VMs are migrated to the destination host which has enough resources for it and where VM consumes equal or less power. This process is repeated for all the VMs and are consolidated into less number of hosts. Hence, by migration of VMs, energy consumption of data center can be reduced.

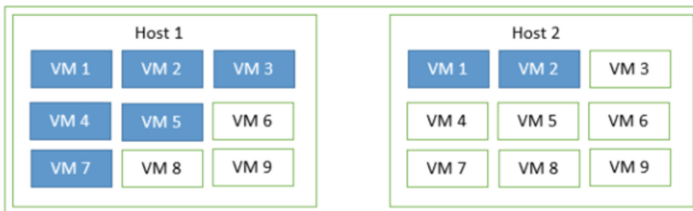


Fig. 2. Snapshot of system without implementation of energy efficient algorithm

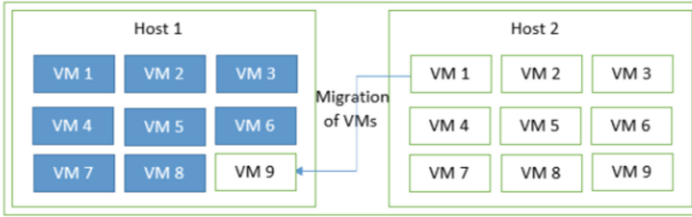


Fig. 3. Snapshot of system after execution of energy efficient algorithm

3.3 Descriptive Statistic: Interquartile Range

Interquartile Range is used to identify over utilized hosts. IQR is chosen, because it is less sensitive to highly skewed data. This method uses historical data of CPU utilization and provides maximum CPU threshold value. It finds the median of data. IQR is difference between median of data series after and before the median value.

$$IQR = m3 - m1$$

$$MaxCPUThreshold = 1 - s * IQR$$

Here, s [8] is a safety parameter, which shows that how aggressive consolidation of VMs are done inside host.

4 Proposed Algorithm

To reduce energy consumption, VMs are needed to be collected from unutilized hosts. For this purpose, first over utilized hosts needs to be identified, IQR is used for it. Host with CPU utilization more than $MaxCPUUtilization$ are over utilized host. For collecting VMs from host three methods are preferred: MMT, RC and Maximum Correlation (MC). In minimum migration time, VMs are selected which requires less time for migration. Migration time is calculated as amount of RAM used by VM divided by spare BW available for selected Host. In Random choice method, VM is selected randomly. In Maximum correlation, VMs are chosen based on maximum correlation of utilization of CPU with other VMs utilization of CPU. Multiple correlation coefficient is used for identifying a correlation between utilization of a CPU of VMs [8]. We use minimum migration time technique for collection of VM. After choosing VM, the host is switched off, if it is empty. Under utilized hosts are collected by excluding over utilized and switched off hosts. All the VMs from such host are collected and consolidate into less number of hosts.

VMs are required to place in appropriate hosts as per polices and availability of resources inside host. VMs are sorted in decreasing order of their utilization of CPU. Power consumption of all the VMs are compared and maximum value is assigned to $minPowerConsumption$. For every VM, hosts are checked for required

Algorithm 1. VM Optimization

```

1: H: Host List,  $h_i$ : Each host inside Host List
2: for each ( $h_i \in H$ ) do
3: if ( $h_i.CPUUtilization > MaxCPUThreshold$ ) then
4:   Add VmsToMigrateFromOverloadedHost ( $h_i$ ) to OverLoadedVmList
5: end if
6: Add OverLoadedVmList to VMMappingList
7: excludeHosts.add(OverUtilizedHosts)
8: excludeHosts.add(SwitchOffHosts)
9: excludeHosts.add(HostsInMigration)
10: end for
11: for each ( $h_i \in H$ ) do
12: if (excludeHosts.contain( $h_i$ )) then
13:   continue()
14: end if
15: if (utilization > 0 AND utilization < minUtilization) then
16:   minUtilization = utilization;
17:   underUtilizedHost = host;
18:   Add vm.underUtilizedHost to VMMappingList
19:   host.deallocateVM()
20: end if
21: end for
22: return VMMappingList

```

Algorithm 2. VM Placement

```

1: H: List of Switched Off and unutilized Host
2: vmList.sortDecreasingUtilization()
3: minPowerConsumption = MAX()
4: for each ( $v_i \in VMMappingList$ ) do
5: allocatedHost = NULL
6: for each ( $h_i \in H$ ) do
7: if ( $h_j.hasResources(v_i)$ ) then
8:   powerReqByVm = estimatePower( $h_j, v_i$ )
9:   if (powerReqByVm < minPowerConsumption) then
10:    allocatedHost =  $h_j$ 
11:    minPowerConsumption = powerReqByVm
12:   end if
13:   if (allocatedHost <> NULL) then
14:    allocationOfHost.add(vm, allocatedHost)
15:   end if
16: end if
17: end for
18: end for
19: return allocation

```

resources. If the host contains sufficient resources for VM, then the power consumption of VM is compared with `minPowerConsumption`, and if it is less, then VM is allocated to that host. This way by analyzing the usage statistics, consolidation of VMs in less number of hosts is achieved. In result, total energy consumed by the data center is reduced, which is shown further in results.

5 Performance Analysis

The proposed technique is implemented on Private Cloud Infrastructure. For evaluation of the results generated after deploying an algorithm on the cloud, sufficient amount of resources are considered. It was very challenging to, because of it's complexity. For performing experiment dual-core HP ProLiant ML110 G4 and HP ProLiant ML110 G5 are used, with MIPS ranging from 1860 to 2660, RAM 4096, storage 10000000 and BW 10000000. VMs characterized with MIPS range from 500 to 2500, RAM range from 613 to 1740, Storage 2500 and BW 100000.

Power Model. By recent study [10,12], in data center CPU has linear power to frequency relationship with DVFS application. It also shows that the server running in idle mode approximately consumes 70% of the energy used by CPU working at full speed. So, to avoid this, the load is consolidated in less number of hosts. The model of power be defined by [6],

$$p(u) = k * P_{max} + (1 - k) * P_{max} * u$$

Here, Pmax: Max power consumption when server is utilized fully, k: part of power consumed when the server is idle mode, u: utilization of CPU. As workload is not stable, Utilization of CPU change over the time. So, it is a function of time, defined as $u(t)$. Total Energy consumption E, defined as [6],

$$E = \int_{t_0}^{t_1} P(u(t))dt$$

Power consumption according to utilization is shown in the Table 1 [8],

Table 1. Power consumption according to CPU utilization [8]

Server	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
HP ProLiant G4	86	89.4	92.6	96	99.5	102	106	108	112	114	117
HP ProLiant G5	93.7	97	101	105	110	116	121	125	129	133	135

5.1 Evaluation of Result

For implementation, 3 physical hosts are taken. Results shows that, the hosts consumes an equal amount of energy during whole time frame if it is ideal or fully utilized. This is results in wastage of resources and increase in energy consumption. By implementing proposed algorithm, resource utilization can be increased and energy consumption can be reduced.

The results after implementation of algorithm shows that, Energy consumption of hosts is decreased by migrating VMs from unutilized hosts and consolidating them in less number of hosts. The host does not consume energy when it is not utilized.

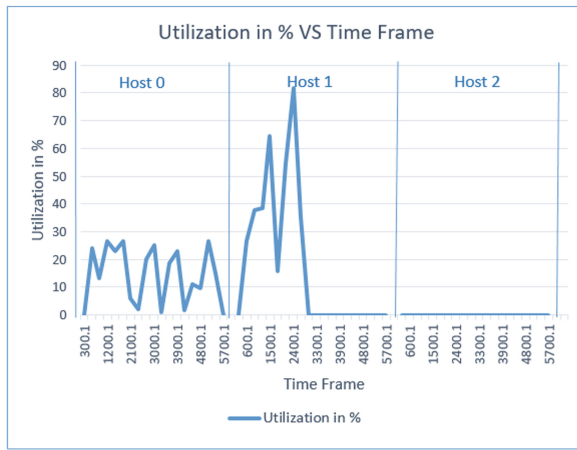


Fig. 4. Time frame VS Utilization without implementation of energy efficient algorithm

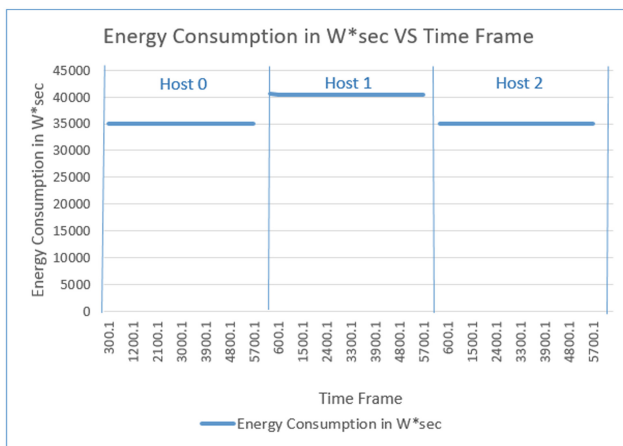


Fig. 5. Time frame VS Energy consumption without implementation of energy efficient algorithm

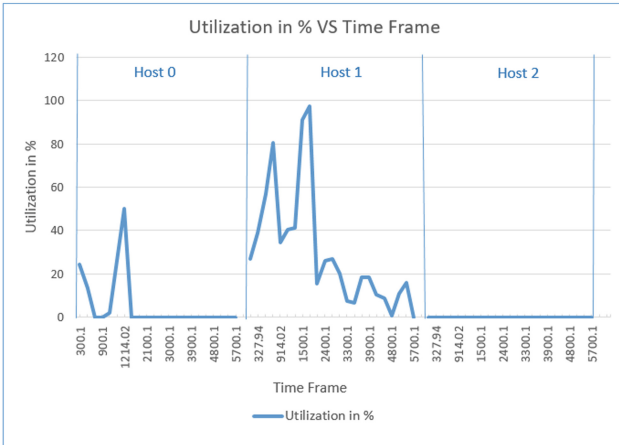


Fig. 6. Time frame VS Utilization after implementation of energy efficient algorithm

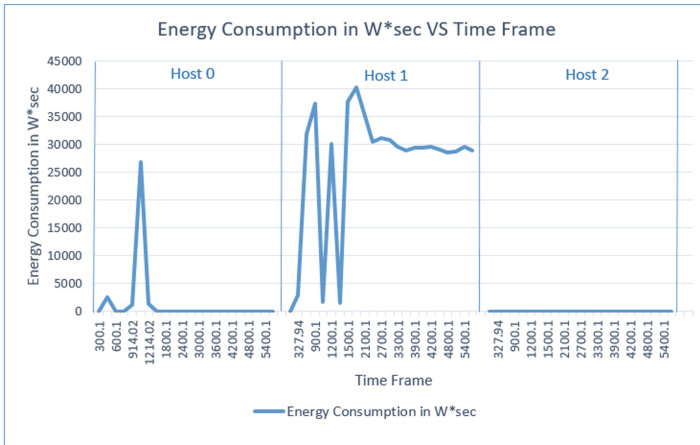


Fig. 7. Time frame VS Energy consumption after implementation of energy efficient algorithm

6 Conclusion and Future Work

Energy Efficient Computing reduces Carbon Footprint in the environment and increases return of investment in cloud resources. Proposed technique performs energy efficient computing at VM level and increases utilization of resources, which in turn reduces active number hosts and energy consumed by the host. This way energy consumed by the data center is decreased.

The research work is intended to analyse energy consumption of host machine in cloud and perform appropriate VM migration in case of overconsumption of

energy. As hardware resources are majorly handled in IaaS, PaaS and SaaS services for energy monitoring are kept as future work.

References

1. Cloud computing and sustainability: The environmental benefits of moving to the cloud (2010)
2. Building return on investment from cloud computing - executive summary (2017)
3. NIST cloud computing program (2017)
4. Akoush, S., Sohan, R., Rice, A., Moore, A.W., Hopper, A.: Predicting the performance of virtual machine migration. In: 2010 IEEE International Symposium on Modeling, Analysis & Simulation of Computer and Telecommunication Systems (MASCOTS), pp. 37–46. IEEE (2010)
5. Beloglazov, A.: Energy-efficient management of virtual machines in data centers for cloud computing. Ph.D. thesis (2013)
6. Beloglazov, A., Abawajy, J., Buyya, R.: Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. *Future Gener. Comput. Syst.* **28**(5), 755–768 (2012)
7. Beloglazov, A., Buyya, R.: Energy efficient allocation of virtual machines in cloud data centers. In: 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing (CCGrid), pp. 577–578. IEEE (2010)
8. Beloglazov, A., Buyya, R.: Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers. *Concurr. Comput. Pract. Exp.* **24**(13), 1397–1420 (2012)
9. Beloglazov, A., Buyya, R., Lee, Y.C., Zomaya, A., et al.: A taxonomy and survey of energy-efficient data centers and cloud computing systems. *Adv. Comput.* **82**(2), 47–111 (2011)
10. Fan, X., Weber, W.-D., Barroso, L.A.: Power provisioning for a warehouse-sized computer. In: ACM SIGARCH Computer Architecture News, vol. 35, pp. 13–23. ACM (2007)
11. Garg, S.K., Buyya, R.: Green cloud computing and environmental sustainability. In: Murugesan, S., Gangadharan, G. (eds.) *Harnessing Green IT: Principles and Practices*, pp. 315–340. Wiley, London (2012)
12. Kusic, D., Kephart, J.O., Hanson, J.E., Kandasamy, N., Jiang, G.: Power and performance management of virtualized computing environments via lookahead control. In: 2008 International Conference on Autonomic Computing, ICAC 2008, pp. 3–12. IEEE (2008)
13. Verma, A., Dasgupta, G., Nayak, T.K., De, P., Kothari, R.: Server workload analysis for power minimization using consolidation. In: *Proceedings of the 2009 Conference on USENIX Annual Technical Conference*, p. 28. USENIX Association (2009)