# A LITERATURE REVIEW: EFFICIENT VM MIGRATION TECHNIQUES FOR ENERGY REDUCTION IN CLOUD COMPUTING

Navneet Singh
Department of Computer Applications,
Sant Baba Bhag Singh University,
Punjab, India.

Dr Vijay Dhir
Director R & D
Sant Baba Bhag Singh University,
Punjab, India.

*Abstract:* Broadly speaking cloud computing is nothing but a highly 'utilitarian' orientation of IT services where users benefited on a pay-as-you go basis. In a way it enables the hosting of pervasive applications from consumer, scientific and business domain. We expect all electronic gadgetry to be 'energy efficient' to possibly achievable limits. So our data centers hosting cloud application must be cost effective and the same time should avoid undue burden of carbon footprint. While excising economy on power consumption by (data center) outmost care needs to be taken so that it never at the cost services provided to end user i.e. SLA violation must be kept as low as possible. Virtualization technology is one of the key features in cloud data centers that can improve the efficiency of hardware utilization through resource sharing, migration, and consolidation of workloads. In this paper we shall cover VM Migration Algorithms for energy reduction in cloud computing along with other novel techniques.

*Keywords:* Cloud Computing, Virtualization, Heuristic, Meta- Heuristic,VM  migration.

## 1. INTRODUCTION

CLOUD computing has revolutionized the Information and Communication Technology (ICT) industry by enabling on-demand provisioning of elastic computing resources on a pay-as-you-go basis. Cloud computing is very much beneficial for small to medium organization. SME ( Small to medium enterprise) can save on up-font cost by outsourcing its computational needs to cloud provider and consequently costs of maintenance and upgrades. Second option is to buid private cloud within organization to boost effective resource managemnet and resource provisioning.

Today large-scale data center contains thousands of computer nodes. These data center consumes huge amount of electric power and $CO_2$ (carbon dioxide) emission to environment. Given the increasing focus on reducing energy use, *ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) recently created Standard 90.4-2016, "Energy Standard for Data Centers."(source: http://www.achrnews.com)*

Internet of things (IoT) is also a rapidly growing branch of IT sector [1]. IoT also makes us of cloud computing. Today large-scale data center contains thousands of computer nodes. These data center consumes huge amount of electric power and $CO_2$ (carbon dioxide) emission to environment. Given the increasing focus on reducing energy use, *ASHRAE (American Society of Heating, Refrigerating and Air-Conditioning Engineers) recently created Standard 90.4-2016, "Energy Standard for Data Centers."(source: http://www.achrnews.com).*

 Eucalyptus cloud software is widely used for creating private cloud for organization[2].
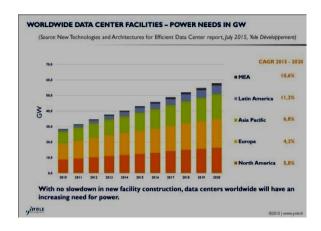
## 2. NEED OF STUDY

The only drawback with cloud computing is that it is a notorious power guzzlers and call for a stringent 'energy efficiency' regime[3]. Currently it is estimated that servers consume 0.5% of the world's total electricity usage. Server energy demand doubles every 4-6 years. With their enormous appetite for energy, today's data centers emit as much carbon dioxide as all of Argentina. Data center emissions are expected to quadruple by 2020. The average data center consumes as much energy as 25,000 households reported by Kaplan et al[4].

Between 2000 and 2007, the total power consumption of datacenters worldwide went from 70 billion to 330 billion kWh; it's projected to grow to more than 1,000 billion kWh by 2020. In 2003, the power density of a single rack of servers was between 250 W and 1.5 kW. In 2014, it had reached almost 10 kW and is projected to reach up to 30 kW by 2020.

"In the actual scenario, with an average Power Usage Efficiency (PUE) of 1.8, worldwide data center energy consumption will reach 507.9 TWh by 2020", explains Mattin Grao Txapartegy, Technology & Market Analyst at Yole."



Source : http://www.yole.fr

## 3. VIRTUAL MACHINE MIGRATION TECHNIQUES

The process of movement of virtual machine instances from one physical node or storage location to another is called VM migration. We are living in an age, where something may be manipulated or altered with the assistance of advance technology [5].

### 3.1 Live Migration and High Availability
Live migration is movement of a virtual machine instance from one physical host to anther while being powered on. Live migration is helpful in load balancing or while performing proactive maintenance in case of failure of physical machine.

### 3.2 Clod/Regular Migration
Cold migration can be defined as migration of powered-off or suspended virtual machine. Cold migration is simple to implement as compared to live migration.
By using cold migration you can also move associated disks from one data store to another.

Before you perform VM migration to reduce energy consumption of a data center following points must be analyzed:
(a)When to migrate VMs[6] Determine the best time to migrate the VM instance to reduce energy consumption without violating SLA.
(b)Which VMs to migrate: once decision has been taken to perform VM migration , next step is select set VM instance for migration from one host to another; that results in effective usage of resources in cloud.
(c) Where to migrate the VMs selected for migration[6]: select the set of physical machines on which VM instances to be migrated.
d)When and which physical nodes to switch on/off: Last step the put the idle nodes to power saver mode or hibernation to reduce energyconsumption.

## 4. CLASSIFICATION OF VM MIGRATION ALGORITHMS FOR ENERGY REDUCTION IN CLOUD COMPUTING

Since efficient VM allocation & migration helps in energy reduction of data center, which is often modeled as bin packing problem and has been proved as NP-hard problem [7]. While excising economy on power consumption by (data center) outmost care needs to be taken so that it never at the cost services provided to end user i.e. SLA violation must be kept as low as possible. Fig 1 shows Classification of VM migration Algorithms.
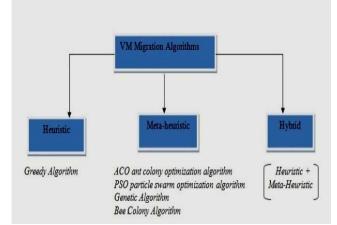


Fig 4 Classification of VM migration Algorithms.

### 4.1 HEURISTIC ALGORITHMS
Heuristic is a set of constraints that aim at finding a good solution for a particular problem [8]. The set of constraints used by heuristic are problem dependent and provide solution to a problem in a limitde time. These heuristic methods have various constraints like number of migrations, SLA, cost, etc. There is need to construct optimization functions in different ways. The main plus point of heuristic algorithms is that they give satisfactory solution to a problem in limited time cost frame. Heuristic algorithms are easier to implement in comparison to meta-heuristic algorithms. Since heuristic algorithms run faster, they are more suitable for online task scheduling that requires minimum response time. Greedy algorithm is a type of heuristic algorithms and is applied in the literature [9][10][11] to quickly obtain a solution for online scheduling scenario.

### 4.2 META-HEURISTIC ALGORITHMS
Meta –heuristic algorithms are mainly designed for a general purpose problem. They follow uniform set  of procedures to construct and solve problems. The typical meta-heuristic algorithms are bio-inspired like genetic algorithms, ACO (Ant colony optimization ,Particle Swarm Optimization  and honey bee foraging algorithms.

### 4.3 HYBRID ALGORITHMS

In Hybrid algorithms, heuristic algorithms are used to provide initial VM placement and meta-heuristic algorithms provides optimum placement of VMs during migration. This algorithm increases the implementation complexity but reduces time and cost space. Thiruvenkadam at el[12] proposed a hybrid genetic algorithm that follows this approach.

## 5. SUMMARY OF TECHNIQUES FOR ENERGY REDUCTION IN CLOUD COMPUTING UNDER QOS CONSTRAINT

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| VRDI (WEN et al.[13]) | Virtual resource dynamic integration (VRDI) method is based on the live migration technology of VM, which can reduce the energy consumption of a data center by integrating the virtual resources. | Simulation with 100 PMs. | Used the encoding, crossover and mutation operations of the genetic algorithm | Synthetic data | Simulation on CloudSim using synthetic data | i) The VRDI method can save about 45% of energy when the resource utilization of PM is less than 50%. (ii)SLA violation of the VRDI method is lower than that of the BMH and ACS-VMC methods. (iii) the number of VMs to be migrated in the VRDI method is larger than the BMH and ACS-VMC methods. |
| ECR (WANG et al. [14]) | New task model that describes the QoS requirements of tasks with the minimum frequency. Used energy consumption ratio (ECR) to evaluate the efficiency of different frequencies under which to execute a take. | three Dell R720 servers with CentOS | DVFS | use NAS Parallel Benchmarks 3.3 (NPB) to generate the testing tasks | real test-bed system and simulation | i) OECR (energy-aware method Optimal Energy Consumption Ratio) can save more than 15% energy as compared to FFD algorithm. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| Self-organized criticality Approach (Laredo et al. [15]) | Self-organized criticality approach for dynamically load-balancing computational workloads. This approach is able to converge towards near-optimal trade-offs where the energetic efficiency and the QoS are maximized. | Simulation with 100 nodes | Bak-Tang-Wiesenfeld sandpile | real trace from the Grid Workloads Archive | Simulation on sandpile simulator and real traces. | Minimizes the energy consumption (i.e., number of active resources through time) and maximizing the QoS (i.e., the average waiting time of tasks in the system |
| Dynamic overbooking strategy (Son et al. [16]) | By using dynamic overbooking strategy, they proposed dynamic overbooking algorithms (ConnCons +MostCorr and ConnDist+Least Corr). | Simulation with 128 hosts | dynamic overbooking strategy | Wikipedia workloads | Simulation on CloudSimSDN | Performs better as compared to baseline algorithms NoOver and ConnNone) in terms of energy consumption and SLA violation. |
| NSGA-II (ZHENG et al. [17]) | Non-dominated sorting genetic algorithm (NSGA-II) and a local selection strategy based on fuzzy is used to generate a Hybrid Energy-Aware Resource Allocation Approach in Cloud Manufacturing Environment | dual 2.63GHz Intel i5 processor and 8 GB RAM. | Genetic & fuzzy | Data collected from 300 high-speed train manufacturers. | Simulation on Matlab 2010b | Minimizes the energy consumption as compared to base line algorithm. |
| Multiagent (MA)-based VM allocation approach (Wang et al. [18]) | MA works by first dispatching a cooperative agent to each PM to assist the PM in managing VM resources efficiently. | Each cloud system consists of 128 PMs.( Tree Network, BCube Network, Lattice-Like Network) | distributed artificial intelligence | Synthetic data | Simulation | Compared with traditional centralized bin packing-based BFD approach, MA performs better in terms of energy & migration cost. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| Novel approach for long-term predictions of resource demands of virtual machines for host overload detection. (Minarolli et al. [19]) | To take into account the uncertainty of long-term predictions, a probability distribution model of the prediction error is built. Based on the probability distribution of the prediction error, a decision-theoretic approach is proposed to make live migration decision that take into account live migration overheads. | Simulation with 100 heterogeneous hosts. | machine learning approach based on Gaussian Processes | PlanetLab workloads | CloudSim simulator | Achieves better performance and higher stability compared to other approaches. |
| MOACS (Ashraf et al. [20]) | Multi-objective ant colony system [MOACS] algorithm for virtual machine consolidation in cloud data centers. | Intel Core i7-4790 Processor with 16 gigabytes of memory. | multi-objective ant colony system | Synthetic data | Simulation using Java program | maximize the number of released PMs thus reduces energy consumption as compared to two existing ant colony optimization based VM consolidation algorithms [Feller-ACO algorithm & single-objective, single-colony ACS VM consolidation algorithm]. |
| Just-in-time adaptive workflow scheduling heuristics (POOLA et al. [21]) | Just-in-time adaptive workflow scheduling heuristics uses on-demand and spot instances to provide fault-tolerant schedules whilst minimizing time and cost. | VMs/cloud resources are modeled similar to Amazon EC2 instances. | ECPTR(*Essential Critical Path Task Replication)* and CTR(Critical Task Replication) | Amazon spot market traces | CloudSim | Cost of ECPTR is much lower than the base line ECPTRRM (*Essential Critical Path Task Replication without Resource Maximization )* algorithm. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| VMCUP-M (Hieu et al. [22]) | virtual machine consolidation algorithm with multiple usage prediction (VMCUP-M) to improve the energy efficiency of cloud data centers | Simulation with 800 heterogeneous hosts | VM consolidation | Real workload traces in the Google Cluster Data dataset & PlanetLab. | CloudSim simulation toolkit | VMCUP-M reduces energy consumption compared to the multiple resource black-box and graybox (BG) scheme. |
| HCLB (Pantazoglou et al. [23]) | HyperCube Load Balancer(HCLB) is a decentralized approach towards scalable and energy-efficient management of virtual machine (VM) instances. | Simulation with 1,024 compute nodes | VM consolidation | Synthetic data | simulation-based implementation in Java | HCLB systematically avoids pushing its compute nodes into the over utilized state, refraining so from such load balancing-induced violations. |
| VM launching overhead reference model (Wu et al. [24]) | The model can accurately predict VM launching overhead within a mean square weighted deviation less than three from all four variables, i.e. VM CPU utilization, system CPU utilization, system I/O utilization and VM launching time | two types of VM instances, i.e., a small instance configured with one virtual CPU core and 2 GB memory, and a large instance configured with 16 virtual CPU cores and 32 GB memory | Cloud bursting | Workload traces from FermiCloud, a private cloud developed by the Fermi National Accelerator Laboratory for scientific workflows. | Simulation with workload traces from FermiCloud | With help reference model efficient resource allocation algorithms can be developed for cloud bursting process to minimize the operational cost and resource waste. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| Brownout (Xu et al. [25]) | brownout approach reduces energy consumption through selectively and dynamically deactivating application optional components. | Simulation with 100 hosts. | VM placement and consolidation algorithm | Workload traces from PlanetLab | CloudSim framework | proposed policies save more energy than the baselines PCO(placement and consolidation algorithm) and UBP(Utilization-based Probabilistic VM consolidation algorithm) |
| M-convex VM consolidation (Huang et al. [26]) | VM consolidation framework based on quasi M-convex optimization framework can achieve a balance among multiple administrative objectives (e.g., power cost, network cost) during the VM consolidation process. | Simulation with 550 servers | VM consolidation | Workload traces from PlanetLab | CloudSim | the proposed framework is efficient, scalable and highly practical. |
| BGM-BLA (Tao et al. [27]) | A binary graph matching-based bucket-code learning algorithm (BGM-BLA) is designed for solving the dynamic migration of VMs (DM-VM ) problem. | Windows 7, 32 bit Core i7-Q720 CPU with 4 GB RAM. | binary graph matching-based bucket-code learning algorithm | Synthetic data | Simulation | BGM-BLA algorithm performs relatively well in terms of the Pareto sets obtained and computational time in comparison with two optimization algorithms, i.e., Non-dominated Sorting Genetic Algorithm (NSGA-II) and binary graph matching-based common-coding algorithm. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| MinES & MinCS (Dai et al. [28]) | two greedy approximation algorithms, minimum energy virtual machine (VM) scheduling algorithm (MinES) and minimum communication virtual machine scheduling algorithm (MinCS), to reduce the energy while satisfying the tenants' service level agreements. | Dell PowerEdge Rack Servers | greedy approximation. | Synthetic workloads and Google's job traces | Simulation & real test-bed. | results demonstrated that MinES and MinCS yield scheduling that are within 4.3 to 6.1 percent energy consumption of the optimal solution while being computationally efficient |
| Platform for virtual machine (VM) placement/migration(VAKILINIA et al. [29]) | Platform for virtual machine (VM) placement/migration is proposed to minimize the total power consumption of cloud data centers (DCs). | Intel Xeon Processor E5-2660 v2 CPUs and 8x16GB DDR3 (M393B2G70DB0-CMA) RAM | (i)estimation module based on ARIMA (Auto Regressive Integrated Moving Average) model was introduced to predict the incoming load of the DC (ii) schedulers were designed to determine the optimal assignment of VMs to the PMs (iii) Column generation (CG) method was applied to solve the large-scale optimization problem. | Synthetic data | MATLAB and IBM ILOG CPLEX | results have shown that the approach explores the optimal solution with an optimality gap of at most 1% in 3 minutes Computation time. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| Green-Works (Li et al. [30]) | Green-Works, a framework for HPC data centers running on a renewable energy mix | IBM System x3650 M2 & Intel Xeon X5570 | Green-Works, a framework | Real-World Workload Traces | Simulation with MatLab | Reduce average performance degradation by only 60% and worst-case performance degradation by 40% while maintaining a desired energy efficiency, battery lifecycle, and backup time. |
| Enhanced bee colony algorithm (Babu et al. [31]) | Enhanced bee colony algorithm for efficient and effective load balancing in cloud environment | Simulation with dynamic environment | bee colony algorithm | Synthetic data | Simulation | reduce the imbalance in the cloud eco system |
| Dynamic VMs placement using PSO (Dashti et al. [32]) | PSO [Particle Swarm Optimization] to guarantee Quality of service of users' tasks, and improve energy efficiency of cloud computing | NA | PSO | Synthetic data | CloudSim | save as much as 14% more energy and the number of migrations |
| Fuzzy logic and heuristic based virtual machine consolidation( Monil et al.[33]) | Fuzzy logic and heuristic based virtual machine consolidation to ensure energy-QoS. balance. | Simulation with 800 heterogeneous physical nodes. | Fuzzy logic and heuristic | work load data provided from CoMon project, a monitoring infrastructure for PlanetLab. | CloudSim | Minimum energy consumption by the proposed method is 102 Kwh therefore we got 8.5 % reduction as compared to baseline algorithm. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| Jettison for partial VM migration (BILA et al. [34]) | An approach that transparently migrates only the working set of an idle VM onto the consolidation server by copying only VM metadata. places desktop PCs in low-power mode when inactive and switches them to running mode when pages are needed by the VM running on the consolidation server | Power Profile of Dell Studio XPS 7100 PC | sleep scheduling and prefetching algorithms | real user traces collected using a Mac OS X-based tracker. | Real & prototype based | Jettison can deliver 44–91% energy savings during idle periods of at least 10 minutes, while providing low migration latencies of about 4 seconds and migrating minimal state that is under an order of magnitude of the VM's memory footprint. |
| Energy-aware resource provisioning framework (Dabbagh et al. [35]) | Framework i) predicts the number of virtual machine (VM) requests, to be arriving at cloud Data centers in the near future, along with the amount of CPU and memory resources associated with each of these requests, ii) provides accurate estimations of the number of physical machines (PMs) that cloud data centers need in order to serve their clients, and iii) reduces energy consumption of cloud data centers by putting to sleep unneeded PMs. | Intel Core 2 Quad Q9400 2.66 GHz processor, and an 8 GB RAM. | modified Best Fit Decreasing (BFD) heuristic for power management | real Google traces | Real & simulation | Proposed energy-aware resource provisioning framework makes substantial energy savings. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| ACS-VMC (Farahnakian et al. [36]) | Ant Colony System (ACS) based VM Consolidation (ACS-VMC) approach to reduces energy consumption while maintaining the required performance levels in a cloud data center | Simulation with 800 heterogeneous PMs | Ant Colony System | Real workload from CoMon project, a monitoring infrastructure for PlanetLab. | Simulation | ACS-VMC consumes less ESV(Energy and SLA Violations) than other benchmarks algorithms in the real workload traces. |
| SmartSLA (Xiong et al. [37]) | a cost-sensitive virtualized resource management system for CPU-bound database services. SmartSLA minimizes the total cost for a DBaaS provider which is composed of SLA penalty cost. | AMD 3.0 GHz Dual cores & Intel Xeon 2.40 GHz Hexa-core | machine learning technologies | web traces from the 1998 World Cup site. | Real environment | SmartSLA is able to minimize the total cost under time-varying workloads compared to the other cost-insensitive approaches. |
| VM workload cycles (Baruchi et al. [38]) | VM workload an important factor and carefully choosing a proper moment to migrate a VM can reduce the live migration penalties. | Simulation | VM migration | Synthetic data | Simulation on cloudsim | Reduce up to 43% of network data transfer and reduce up to 74% of live migration time when compared to traditional consolidation strategies |
| HGA for the VM placement (Tang et al. [39]) | HGA( Hybrid genetic algorithm )for a new virtual machine placement problem that considered the energy consumption in both physical machines and the communication network in a data center | Intel Core 2 Duo CPU of 3.00GHz and a 4.00GB RAM. | Hybrid Genetic Algorithm | Synthetic data | Simulation using java | Hybrid genetic algorithm significantly outperforms the original genetic algorithm, and that the hybrid genetic algorithm is scalable. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| Energy aware profiling(Alzamil et al. [40]) | Energy aware profiling helps us to better understand the way the energy is consumed and help the application developer from the application level enhance their decision-making in terms of energy-awareness when optimizing their applications and services. | Four Dell commodity servers(server consists of a four core X3430 Intel Xeon CPU, running at the default clock speed of 2.40GHz and a total of 8GB of RAM ) | Energy aware profiling | Leeds Testbed | Real & simulation | energy-awareness at physical host and virtual machine levels. |
| PRS(Chen et al. [41]) | PRS (Proactive and Reactive Scheduling)dynamically exploits proactive and reactive scheduling methods, for scheduling real-time, a periodic, independent tasks. To improve energy efficiency, proposed three strategies to scale up and down the system's computing resources according to workload to improve resource utilization and to reduce energy consumption for the cloud data center. | KVM cluster with five Dell Optiplex 7010MT, each of physical host has a CPU (i3 3.9GHz 4cores), 3.7G memory | interval number theory | Synthetic workloads and Google workload traces. | Real environment & simulation with cloudsim | PRS performs better than baseline algorithms, and can effectively improve the performance of a cloud data center. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| Dynamic Resource Allocation using Virtual Machines (Xiao et al. [42]) | The concept of "skewness" to measure the unevenness in the multi-dimensional resource utilization of a server. By minimizing skewness, combine different types of workloads nicely and improve the overall utilization of server resources. develop a set of heuristics that prevent overload in the system effectively while saving energy | 30 Dell PowerEdge blade servers with Intel E5620 CPU and 24GB of RAM. | skewness | traces from servers and desktop computers in university | simulation | Achieves both overload avoidance and green computing for systems with multi-resource constraints. |
| MHOD-OPT (Baloglazov et al. [43]) | Optimal Markov Host Overload Detection (MHOD-OPT) algorithm for the problem of host overload detection as a part of dynamic VM consolidation. | simulation | Markov chain model | PlanetLab Workload Traces | Simulation with cloudsim | Proved best among baseline techniques |
| Novel adaptive heuristics for dynamic consolidation of VMs (Beloglazov et al. [44]) | novel adaptive heuristics for dynamic onsolidation of VMs based on an analysis of historical data from the resource usage by VMs. | Simulation with 800 heterogeneous physical nodes | adaptive heuristics | Real workload from CoMon project, a monitoring infrastructure for PlanetLab | CloudSim toolkit | significantly reduced energy consumption, while ensuring a high level of adherence to the Service Level Agreements |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| VM placement (Wu et al. [45]) | genetic algorithm for a new virtual machine placement problem that considers the energy consumption in both the servers and the communication network in the data center. | Simulation | genetic algorithm | synthetic workloads | GA has been implemented in Java. | On average the solutions produced by the GA are 3:5%-23:5% better than those produced by the FFD(First Fit Decreasing). |
| DCeP (Sego et al. [46]) | Data Center Energy Productivity [DCeP] metric, which is the ratio of useful work produced by the data center to the energy, consumed performing that work. | Pacific Northwest National Laboratory (PNNL) | Data Center Energy Productivity | HPC Workload | Real environment | DCeP can be used to clearly measure and distinguish the energy productivity of different operational states in a data center. |
| Barely alive states (Anagnostopo ulou et al. [47]) | Barely alive states through which we can perform energy efficient cloud computing a family of *barely alive* active low-power server states that facilitates both fast reactivation and access to memory while in a low-power state. | Simulation with 32-node clusters | Barely alive states | Traces from Ask.com | Simulation | Barely alive states can reduce service energy consumption by up to 38%, compared to an energy-oblivious system. |
| CADE(Kaplan et al. [48]) | CADE [Corporate Average Datacenter Efficiency], a new industry standard efficiency measure developed by McKinsey & Company. | NA | CADE | NA | NA | CADE helps in calculating energy efficiency of a data center. |
| VPM (Nathuji et al. [49]) | VirtualPower Management (VPM) approach to online power management | multiple dual core Pentium 4 ma-chines | virtualization | Traces from Delta Air Lines | Real test-bed | VPM provide up to 34% improvements in power consumption. |

| Scheme( Ref.) | Description | Experimental Configuration | Technique used | workload | Evaluation | Performance improvement |
|---|---|---|---|---|---|---|
| Thermal load balancing (Sharma et al. [50]) | local and regional policies for thermal control for reducing energy consumption in data center | data center with with standard racks, each containing 20, 2U (90mm) high, Hewlett-Packard A-Class servers | thermal load balancing | Real traces | fault injection simulations | energy consumption can be reduced by more than 14% by workload placement |
| Energy efficient server cluster(Elnoza hy et al. [51]) | five Server power management technique, these are Independent voltage scaling [IVS], Coordinated voltage scaling [CVS], Vary-On Vary-Off [VOVO], Combined policy & Coordinated Policy to perform energy efficient cloud computing | Simulation | Server power management techniques | Nagano Winter Olympics Servers at Columbus . | Simulation | IVS can only account for energy saving of 20-30%. VOVO-CVS combine help to save up to 18% more energy than VOVO alone but certainly not without complicated implementatio n. |

## 6. CONCLUSION AND FUTURE DIRECTIONS

Multi tenancy, concurrency and distribution are main feature of any cloud computing architecture [52]. This paper covers different Vm Migration Algorithms for energy reduction in cloud computing (Heuristic ,Meta-heuristic& Hybrid algorithms) along with other novel techniques. The main plus point of heuristic algorithms is that they give satisfactory solution to a problem in limited time cost frame. Heuristic algorithms are easier to implement in comparison to meta-heuristic algorithms. In Hybrid algorithms, heuristic algorithms are used to provide initial VM placement and meta-heuristic algorithms provides optimum placement of VMs during migration. This algorithm increases the implementation complexity but reduces time and cost space. Depending upon the kind of problem, you can use one or other. Now, we discuss the future directions and challenges as below:

- We observed that most meta-heuristic s gives better results in terms in energy consumption and QoS(Quality-of-service) but they are tested using only simulators. There is need to test these Meta-heuristic algorithms like genetic algorithms, ACO ( Ant colony optimization ,Particle Swarm Optimization and honey bee foraging algorithms in realistic environment. So that cloud vendors can adopt these methods to reduce energy consumption & ensure QoS of user's tasks.

- We also have observed that how it is difficult to ensure

a balance between energy consumption and SLA violations. Therefore reducing energy consumption while maintain a QoS (Quality-of-Service) is a future research challenge.

- While considering the diversity of our surveyed papers, we want to know which algorithm is best under which environment. This question still is open because of heterogeneity of different algorithms and lack of validation of these algorithms under different realistic environment. A thorough testing of all these algorithms under same environment and under heterogeneous environment is definitely required as future work.

## 7. ACKNOWLEDGMENT

## REFERENCES

[1] Dhir V. et. al., Internet of things: Vision, Challenges and Future Scope, International Journal of Advanced Research in Computer Science, May 2017.

[2] Singh N., Dhir V., Building Private Cloud Using Eucalyptus, Proceeding of 1st National Conference on Emerging Trends on Engineering & Technology (ETET- 2017) at SBBS University, Punjab. International Journal of Advanced

Research in Computer Science, Vol. 8, No. 4, May 2017.

[3] Sigh N., Bal J.S., Energy Efficient Data Center: A systematic literature review, Advances in Computational Science and Technology, Research India Publications, Vol.10, Number 1, pp. 121-128, 2017, ISSN 0973-6107.

[4]. J. M. Kaplan, W. Forrest, and N. Kindler, "Revolutionizing data center energy efficiency,"Technical Report, available at http://www.mckinsey.com-/clientservice-/bto/pointofview/pdf/-revolutionizing data center efficiency.pdf. Accessed on 25-06-2017.

[5] Dhir V. et. al., A Review on Image Forgery & its Detection Procedure, International Journal of Advanced Research in Computer Science, May 2017.

[6]. Baloglazov A., (Thesis) Energy-Efficient anagement of Virtual Machines in Data Centers for Cloud Computing, THE UNIVERSITY OF MELBOURNE, 2013

[7]. Coffman Jr EG, Garey MR, Johnson DS. Approximation algorithms for bin packing: A survey. Approximation Algorithms for NP-Hard Problems. Boston, MA, USA: PWS Publishing Co.; 1996:46–93.

[8]. Talbi EG. Metaheuristics: From Design to Implementation, vol. 74: Hoboken, New Jersey, United States: John Wiley & Sons; 2009.

[9]. Song X, Ma Y, Teng D. A load balancing scheme using federate migration based on virtual machines for cloud simulations. Math Prob Eng. 2015;2015:1–

[10]. Ni J, Huang Y, Luan Z, Zhang J, Qian D. Virtual machine mapping policy based on load balancing in private cloud environment. 2011 International Conference on Cloud and Service Computing (CSC). IEEE; 2011:292–295.

[11]. Tian W, Zhao Y, Zhong Y, Xu M, Jing C. A dynamic and integrated load-balancing scheduling algorithm for cloud datacenters. 2011 IEEE International Conference on Cloud Computing and Intelligence Systems. Beijing, China: IEEE; 2011:311–315.

[12]. Thiruvenkadam T, Kamalakkannan P. Energy efficient multi dimensional host load aware algorithm for virtual machine placement and optimization in cloud environment. Indian J Sci.

[13] WEN Y., LI Z., JIN S., LIN C., LIU Z., Energy-Efficient Virtual Resource Dynamic Integration Method in Cloud Computing, IEEE Special Section On Emerging Trends, Issues, And Challenges In Energy-Efficient Cloud Computing, June 29, 2017.

[14] WANG S, QIAN Z , YUAN J, YOU I, A DVFS Based Energy-Efficient Tasks Scheduling in a Data Center, IEEE Special Section On Emerging Trends, Issues, And Challenges In Energy-Efficient Cloud Computing, July 11, 2017,

[15] Laredo J.L.J., Guinand F.,Olivier D., Bouvry P., Load Balancing at the Edge of Chaos: How Self-Organized Criticality Can Lead to Energy-Efficient Computing, IEEE Transactions On Parallel And Distributed Systems, Vol. 28, No. 2, February 2017.

[16] Son J, Dastjerdi AV, Calheiros RN, Buyya R, SLA-Aware and Energy-Efficient Dynamic Overbooking in SDN-Based Cloud Data Centers, IEEE Transactions On Sustainable Computing, Vol. 2, No. 2, April-June 2017.

[17] ZHENG. H., FENG Y., TAN J., A Hybrid Energy-Aware Resource Allocation Approach in Cloud Manufacturing Environment, IEEE Special Section On Emerging Cloud-Based Wireless Communications And Networks, June 2017.

[18] Wang et al, Multiagent-Based Resource Allocation for Energy Minimization in Cloud Computing Systems, IEEE Transactions On Systems, Man, And Cybernetics: Systems, VOL. 47, NO. 2, February 2017.

[19] Minarolli D., Mazrekaj A., Freisleben B., Tracking uncertainty in long-term predictions for host overload and underload detection in cloud computing. Journal of

Cloud Computing Advances, Systems and Applications, Springer, 2017.

[20]. Ashraf A., Porres I., Multi-objective dynamic virtual machine consolidation in the cloud using ant colony system, International Journal of Parallel, Emergent and Distributed Systems, Taylor &Francis, 2017.

[21] POOLA D., RAMAMOHANARAO K., BUYYA R., Enhancing Reliability of Workflow Execution Using Task Replication and Spot Instances, ACM Transactions on Autonomous and Adaptive Systems, Vol. 10, No. 4, Article 30,February 2016.

[22] Hieu N. T., Francesco M. D, Yl a-Jaaski A., Virtual Machine Consolidation with Multiple Usage Prediction for Energy-Efficient Cloud Data Centers, IEEE Transaction On Services Computing, March 2016.

[23] Pantazoglou M., Tzortzakis G., Delis A., Decentralized and Energy-Efficient Workload Management in Enterprise Clouds, IEEE Transactions On Cloud Computing, Vol. 4, No. 2, April-June 2016.

[24] Wu H, Ren S, Garzoglio G, Timm S., Bernabeu G., Chadwick K.,Noh S-Y, A Reference Model for Virtual Machine Launching Overhead, IEEE Transactions On Cloud Computing, Vol. 4, No. 3, July-September 2016.

[25] Xu M., Dastjerdi A.V., Buyya R., Energy Efficient Scheduling of Cloud Application Components with Brownout, IEEE Transactions On Sustainable Computing, Vol. 1, No. 2, July-December 2016.

[26] Huang Z. et al, M-Convex VM Consolidation: Towards a Better VM Workload consolidation, IEEE Transactions On Cloud Computing, Vol. 4, No. 4, October- December 2016.

[27] Tao F. et al, BGM-BLA: A New Algorithm for Dynamic Migration of Virtual Machines in Cloud Computing, IEEE Transactions On Services Computing, Vol. 9, No. 6, November/December 2016.

[28] Dai X. et al, Energy-Efficient Virtual Machines Scheduling in Multi-Tenant Data Centers, IEEE Transactions On Cloud Computing, VOL. 4, NO. 2, APRIL-JUNE 2016.

[29] VAKILINIA S. et al , Energy Efficient Resource Allocation in Cloud Computing Environments, IEEE Special Section On Future Networks: Architectures, Protocols, And Applications, December 1, 2016.

[30] LI C., Wang R., Qian D., Li T., Managing Server Clusters on Renewable Energy Mix. ACM Transactions on Autonomous and Adaptive Systems, ACM, 2016.

[31]. Babu K.R.R., Samuel P, Enhanced Bee Colony Algorithm for Efficient Load Balancing and Scheduling in Cloud, Innovations in Bio-Inspired Computing and Applications, vol 424, pp 67-78, springer, 2016

[32] Dashti S.E., Rahmani A.M., Dynamic VMs placement for energy efficiency by PSO in cloud computing. Journal of Experimental & Theoretical Artificial Intelligence, [Taylor &Francis, 2016].

[33] Monil A. H., Rahman R. M., VM consolidation approach based on heuristics, fuzzy logic, and migration control. Journal of Cloud Computing Advances, Systems and Applications, Springer, 2016.

[34] BILA N., WRIGHT E.J., LARA E.D., JOSHI K., LAGAR-CAVILLA H.A., PARK E., GOEL A., TUNEN M.H., SATYANARAYANAN M., Energy-Oriented Partial Desktop Virtual Machine Migration, ACM Transactions on Computer Systems, Vol. 33, No. 1, Article 2, Publication date: March 2015.

[35] Dabbagh M., Rayes A. , Energy-Efficient Resource Allocation and Provisioning Framework for Cloud Data Centers, IEEE Transactions on Network and Service Management , December 2015.

[36] Farahnakian F. et al., ``Using ant colony system to consolidate VMs for green cloud computing,'' IEEE Trans. Serv. Comput., vol. 8, no. 2, pp. 187198, Apr. 2015.

[37] Xiong P. et al, SmartSLA: Cost-Sensitive Management of

Virtualized Resources for CPU-Bound Database Services, IEEE Transactions On Parallel And Distributed Systems, VOL. 26, NO. 5, MAY 2015.

[38] Baruchi A. et al, Reducing Virtual Machine Live Migration Overhead via Workload Analysis, IEEE Latin America Transactions, VOL. 13, NO. 4, APRIL 2015.

[39] Tang M., Pan S., A Hybrid Genetic Algorithm for the Energy-efficient Virtual Machine Placement Problem in Data Centers, Neural Processing Letters, springer, 2015.

[40] Alzamil I., Djemame K., Arstrong D., Kavanagh R., Energy-Aware Profiling for Cloud Computing Environments. Electronic Notes in Theoretical Computer Science, Elsevier, 2015.

[41] Chen H., Zhu X., Guo H., Zhu J., Qin X., Wu J., Towards Energy-Efficient Scheduling for Real-time tasks under uncertain Cloud Computing, Journal of System and Software Vol. 99, pp. 20-35, Elsevier, 2015.

[42] Xiao Z., Song W., Chen Q., Dynamic Resource Allocation using Virtual Machines for Cloud Computing Environment. IEEE Transaction on Parallel and Distributed Systems, Vol. 24, No. 6, 2013.

[43] Baloglazov A., Buyya R., Managing Overloaded Hosts for Dynamic Consolidation of Virtual Machines in Cloud Data Centers Under Quality of Service Constraints. IEEE Transactions on Parallel and Distributed System, 2013, pp. 1366-1379.

[44] 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, CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE, Wiley InterScience, 2012.

[45] Wu G., Tang M., Tian Y.C., Li W., Energy-efficient Virtual Machine Placement Problem in Data Centers by Genetic Algorithms, Neural Processing Letters, springer, 2012.

[46] Sego L. H., Marquez A., Rawson A., Cader T., Fox K., Gustafson W. I., Mundy C.J., Implementing the Data Center Energy Productivity Metric. ACM Journal on Emerging Technologies in Computing Systems, 2012.

[47] Anagnostopoulou V., Biswas S., Saadeldeen H., Savage A., Bianchini R, Yang T., Franklin D., and Chong F. T., "Barely alive memory servers: Keeping data active in a low-power state," ACM Journal on Emerging Technologies in Computing Systems [JETC], vol. 8, no. 4, pp. 31:1–31:20, 2012.

[48] Kaplan J.M., Forrest W., Kinder Noah, Revolutionizing Data Center Energy Efficiency, Mckinsey & Company, july 2008.

[49] Nathuji R., Schwan K., Virtual power: coordinated power management in virtualized enterprise systems, ACM SIGOPS Operating Systems Review 41 [6] , pp. 265–278, 2007.

[50]. Sharma R.K , Bash C.E, Patel C.D, Freidrich R.J, Chase J.S., Balance of power: Dynamic thermal management for internet data centers. IEEE Internet Computing [Volume:9, Issue: 1], pp. 42-49, IEEE, 2005.

[51] Elnozahy E.N.M., Kistler M., Rajamony R., Energy-efficient server clusters. Power-Aware Computer Systems, Springer, 2002.

[52] Dhir V. et. al., Architectural and Qos Issues in Mobile Cloud Computing Environment for Real-Time Video Streaming, International Journal of Advanced Computer Science and Applications, Vol. 7, No. 1, 2016.