

# Pervasive SLA and Energy Aware Dynamic Virtual Machines Consolidation in Cloud Data Centers

Priya Rana

M.Tech. Scholar

Department of Computer Science  
& Engineering

Sagar Institute of Science  
Technology & Research, Bhopal

Amit Ganguli

Associate Professor

Department of Computer Science  
& Engineering

Sagar Institute of Science  
Technology & Research, Bhopal

Preeti Pathariya

M.Tech. Scholar

Department of Computer Science  
& Engineering

Sagar Institute of Science  
Technology & Research, Bhopal

## ABSTRACT

The Cloud Computing (CC) model is also referred to as Pervasive Computing and proved promising by complicated automation, provisioning and virtualization technologies. The shifts to the computational demands results in greater power consumption, increased operational costs and high carbon emissions to environment. The challenge for the Cloud Provider is to deal with necessary requirement of power-performance trade-off by satisfying high Quality of Service (QoS) defined by Service Level Agreements (SLAs) while maximizing their profits. Out of several issues, Optimization of Energy consumption has gain extensive attention for enhancing the profit. Dynamic Virtual Machine (VM) Consolidation is potential approach for reducing energy consumption by dynamically adjusting the number of active machines to match resource demands and it is one of the most important challenges in the ubiquitous computing. The theme of this work is to propose the 'Pervasive SLA and Energy Aware Dynamic VM Consolidation' policy and provide the baseline for better performance and environment. By conducting a performance evaluation studies a comparative analysis of proposed and various existing energy efficient VM consolidation techniques are presented. For experimentation purpose, in CloudSim toolkit, real world workload traces from more than a thousand VMs are taken. The results help in analyzing the effectiveness of existing policies. The experimental results also demonstrates that the proposed policy is scalable and offers substantial cost savings by saving energy while effectively dealing with firm QoS requirements negotiated by SLA.

## Keywords

Cloud Computing, CloudSim, Data Center, Energy Aware, Pervasive Computing, Quality of Service, Service Level Agreement, Ubiquitous Computing Virtualization, Virtual Machine, Virtual Machine Consolidation,

## 1. INTRODUCTION

The growth of Cloud Computing (CC) model has resulted in the establishment of huge data centers (DCs) containing thousands of computing nodes and cooling systems thus consuming enormous amounts of electrical energy. These DCs hosting cloud applications consume huge amounts of energy, leading to high operational cost and greenhouse gas emission [1]. Some estimation mentioned that by 2014 infrastructure and energy costs would contribute about (3/4)th whereas cost of IT would be just (1/4)th to the overall cost of operating a DC [2]. The inefficient utilization of servers from the energy consumption perspective can be handled by using the capabilities of latest virtualization technology [3, 4]. This allows Cloud providers (CP) to create multiple Virtual

Machine (VM) instances on a physical server, consequently improving the utilization of resources at DC. Other way to preserve energy in cloud DCs is to adjust idle servers into a power saving state during periods of low utilization. Another emerging and efficient solution to conserve energy is Dynamic VM Consolidation which suggests repacking of VMs on the least number of physical machines (PMs).

Dynamic VM consolidation includes efficient determination of overload and under-loaded hosts, so that effective decisions - like migration of VMs to other hosts, can be taken to optimize the power. The migrated VMs dynamically consolidated to the least number of physical nodes in accordance with the current resource requirements [5]. In this work, Energy-Efficient VM Consolidation Policies which meets QoS expectations are explored to propose a 'Pervasive SLA and Energy Aware policy for Dynamic VM Consolidation in Cloud DCs.

The paper outlines the literature review in section 2. The proposed work is presented in section 3. Experimental Setup and Result Analysis is depicted in section 4. Finally the paper is concluded with future directions in section 5.

## 2. LITERATURE REVIEW

With an enormous growth of virtual computing environments such as Clouds the context has been shifted to DCs. A description of different techniques and approaches for reducing energy consumption is surveyed in [6]. Current state-of-the-art Cloud infrastructure such as Amazon EC2 [7] neither support energy-efficient resource allocation that considers consumer preference for energy saving schemes, nor utilize sophisticated economic models to set the right incentives for consumers to reveal information about their service demand accurately [8]. Consequently, providers cannot accomplish efficient service allocation, which meets consumer requirements and expectations with regard to their energy saving goal for Green CC. Starting with most common Dynamic Voltage and Frequency Scaling (DVFS) [9, 10] technique, survey of various papers is presented below:

DVFS [9, 10] technique allows the dynamic adjustment of the voltage and frequency of the CPU based on the current resource requirements. Most modern CPUs in mobiles, desktops, and server systems support DVFS. Energy consumption can also be reduced by applying DVFS to all the CPUs in data centers. The continuous improvements of the CPU power efficiency are in effort with enablement of power-saving techniques (e.g., DVFS) with OS.

The authors [11] propose an admission control and scheduling mechanism which maximize the resource utilization, profit and also ensure QoS requirements of users to met specified

SLAs. Authors assumed two types of application workloads having different QoS requirements, i.e., transactional and non-interactive batch jobs.

Anton Beloglazov et al. [12] presented a survey of research in energy-efficient computing. A number of open research challenges addressed includes: The architectural principles for energy-efficient management of Clouds; energy-efficient resource allocation policies and scheduling algorithms considering QoS expectations and power usage characteristics of the devices; Both resource providers and consumers are substantially benefited.

Beloglazov and Buyya [13, 14] have proposed a novel technique for dynamic consolidation of VMs based on accustomed (adaptive) utilization thresholds, which ensures a high level of meeting the Service Level Agreements (SLA).

Jung et al. [15] have investigated the problem of dynamic consolidation of VMs running a multi-tier web-application using live migration, while meeting SLA requirements.

The authors [16] applied weighted linear regression to envisage the probable workload and passively optimize the resource allocation to implement an energy-aware dynamic VM consolidation framework.

VMs. Bobroff et al. [7] proposed a forecasting technique to determine server overload. The forecasting is based on time-series analysis of historical data.

Verma et al. [8] represented the utilization of the CPU by a VM as a uniformly distributed random variable. The problem of dynamic placement is formulated as continuous optimization in heterogeneous environment. The VM placement is optimized to lessen the power consumption and enhance the performance at each time frame.

Adaptive policies for Dynamic VM Consolidation presented in [17] are studied and briefly described. Various heuristics have been proposed in the literature for deciding the migration time of VMs, from a host based on utilization thresholds. These techniques will provide seamless benefits to environment and increase CP experience and reduce energy consumption at DCs. **The work proposed in this paper is compared with the relevant techniques presented in [17].**

### 3. PROPOSED WORK

#### 3.1 Architecture of Proposed work

Figure 1 shows the high-level architecture for supporting energy-efficient service allocation in a Green Cloud computing infrastructure [17, 18]. The targeted system for purpose of comparing and proposing VM consolidation policy, considered in this work is an IaaS environment represented by a large-scale DC consisting of M heterogeneous physical nodes.

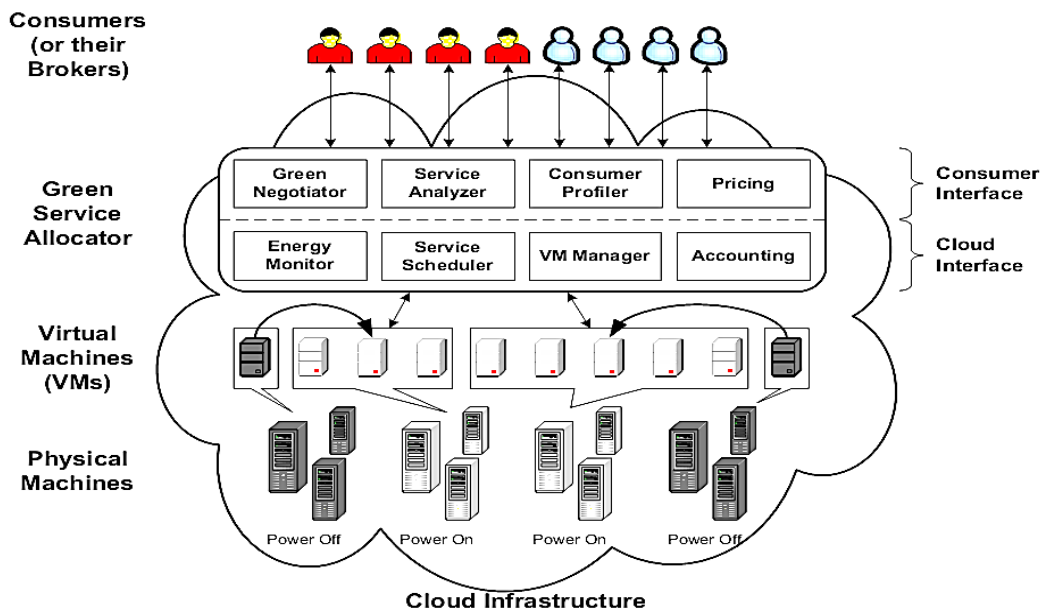


Figure 1: System Model Chosen for Proposed Work [18]

The power model and energy consumption model as described in [13, 18], is utilized in this work. The Power model defines the power consumption as a function of the CPU utilization  $P(u)$ , as shown in equation (1) and total energy consumption by a server as defined in equation (2), where  $P_{max}$  is the maximum power consumed;  $k$  is the fraction of power consumed by an idle server; and  $u$  is the CPU utilization.

$$P(u) = k.P_{max} + (1 - k).P_{max}.u = P_{max}.(0.7 + 0.3.u) \quad (\text{Equation 1})$$

$$E = \int_t P(u(t))dt \quad (\text{Equation 2})$$

The three processes of general algorithm for Dynamic VM consolidation framework involves: i. detection of overloaded and under-load hosts. ii. Selection of VMs to be migrated. List of VMs is prepared that are selected to be migrated. iii. Placement of VMs from these hosts is also determined. VM placement algorithm is called to find a new placement for the VMs to be migrated.

#### 3.2 Proposed Host Overload and Under-load Detection Algorithm: “Effective Utilization Algorithm”

The proposed Overload and Under-Load Detection Algorithm is named as “Effective Utilization” and represented as “Efu”. In this algorithm we adapted a pervasive heuristic of

maximum and minimum utilization as performance of system which is also based on maximum demand by workloads. When the process of detecting overload or under-load detection is invoked it compares the current CPU utilization with dynamic heuristic of all the hosts with the defined threshold. Based on this dynamic heuristic of maximum and minimum utilization threshold, the algorithm detects a host overload or under-load. Then the process tries to place the VMs, selected for migration from this host on other hosts maintaining them not been overloaded or under-loaded.

### 3.3 Proposed VM Selection Algorithm: “Least Migration Time”

At the next step particular VMs are selected to migrate from that host. VM selection algorithm will be iteratively applied until the host is considered as not being overloaded. Here, the proposed policy for VM selection is named as “Least Migration Time” and represented as “Lmt”. The proposed “Least Migration Time” policy, considers two pervasive metrics: RAM utilization and current requested total CPU MIPS. Such metrics are chosen to optimize the selection of

such VMs that are utilizing fewer resources on the present host and of course take less migration time.

### 3.4 VM Placement

For VM placement it is reasonable to apply a heuristic, such as the Power Aware Best Fit Decreasing (PABFD) algorithm [18], which provided the optimal solution.

## 4. EXPERIMENTAL SETUP

It is really difficult to conduct exhaustive large-scale experiments on a real cloud which is required to evaluate and compare the algorithms discussed in the literature. So, simulations have been done on CloudSim [14, 19, 20] toolkit fronted by Buyya, which is a modern simulation framework aimed at CC environments. This toolkit is extensible and offers simulation of various VM consolidation policies proposed in [17].

### 4.1 Simulation Scenario

The DC configurations and general parameters that are considered for simulation purposes are mentioned in Table-1

**Table 1: DC Configuration and Simulation Parameters**

Variables	Value	Remark
<b>Datacenter</b>	Architecture X86 Operating System: Linux VM Manager: Xen	
<b>Planet Lab workload</b>	20110303	
<b>NUMBER_OF_HOSTS</b>	800 of 2 types	2 CPU cores each.
<b>No. of VMs</b>	1052; Types 04	Each 1 CPU core with maximum of 2500, 2000, 1000, 500 MIPS, 1 GB RAM, 1 GB/s network bandwidth and 1 GB of storage.
<b>VM Allocation Policy</b>	Effective Utilization (efu)	Thr, Mad, Iqr [17]
<b>VM Selection Policy</b>	Less Migration Time (lmt)	Mmt [17] and Mu [17]
<b>RAM utilization threshold</b>	0.90	
<b>CPU utilization threshold</b>	0.90	
<b>SCHEDULING_INTERVAL</b>	900 s	15 min.
<b>SIMULATION_LIMIT</b>	24 * 60 * 60;	24 hrs.
<b>Power Dataset</b>	SPECpower	Available at [21, 22]
<b>Workload</b>	Randomly assigned to each VM	
<b>Cloudlets / Tasks</b>	2500 * Simulation Limit	Full size: 300MB Output size: 300 MB

## 4.2 Implementation

The implementation absorbs modification to several java files under Cloudsim toolkit. Some of the java files (e.g.

Constants.java and PlanetLabConstants.java) involving configuration parameters are taken as they are. Figure 2 shows snapshot of one of the obtained simulation results.

Figure 2: Screenshot of Results

## 4.3 Result, Evaluation and Analysis

The evaluation is done for few important performance parameters mentioned in [16] are presented below:

- **Total energy consumption**

Total energy consumption is defined as the sum of energy consumed by the physical resources of a DC as a result running application workloads.

- **SLA violation**

When a VM cannot get the promised Quality of Service (QoS), SLA violation takes place. For example when a VM cannot get requested MIPS, SLA violation issue occur.

The comparison between Dvfs, Thr-Mmt, Mad-Mmt and Iqr-Mmt on basis of 2 parameters is shown in Table 2. The policies Thr-Mmt, Mad-Mmt, Iqr-Mmt are proposed in [17].

Table 2: Comparison of DVFS, IqrMmt, MadMmt, ThrMmt and EfuLmt

Parameters	DVFS	IqrMmt	MadMmt	ThrMmt	Proposed EfuLmt
Energy Consumption (kWh)	817.60	213.31	212.30	225.28	210.05
Average SLA violation (%)	0	10.08	10.19	10.11	10.16

The comparison between Dvfs, ThrMu, MadMu, and IqrMu on basis of five parameters is shown in Table 3. The policies Thr-Mu, Mad-Mu, Iqr-Mu are proposed in [17].

Table 3: Comparison of Dvfs, IqrMu, MadMu, ThrMu and EfuLmt

Parameters	Dvfs	IqrMu	MadMu	ThrMu	Proposed EfuLmt
Energy Consumption (kWh)	817.6	213.31	212.30	217.34	210.05
Average SLA violation (%)	0	10.08	10.19	10.20	10.16

The comparative graph of Energy consumption and Average SLA violation of each of above mentioned schemes and

proposed scheme are represented in Figure 3 and 4 respectively.

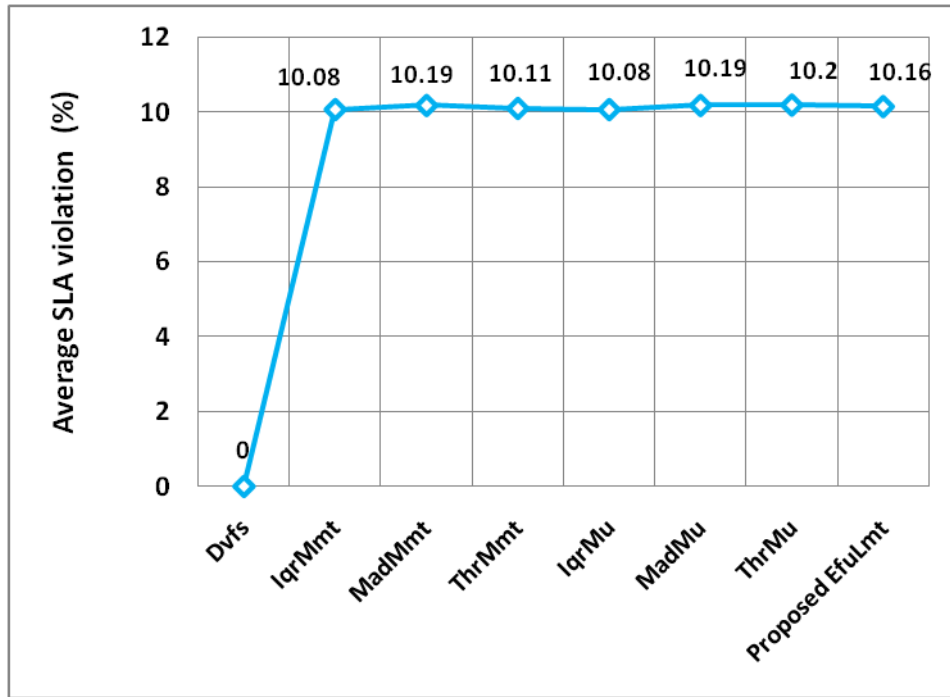


Figure 3: Average SLA violations (%)

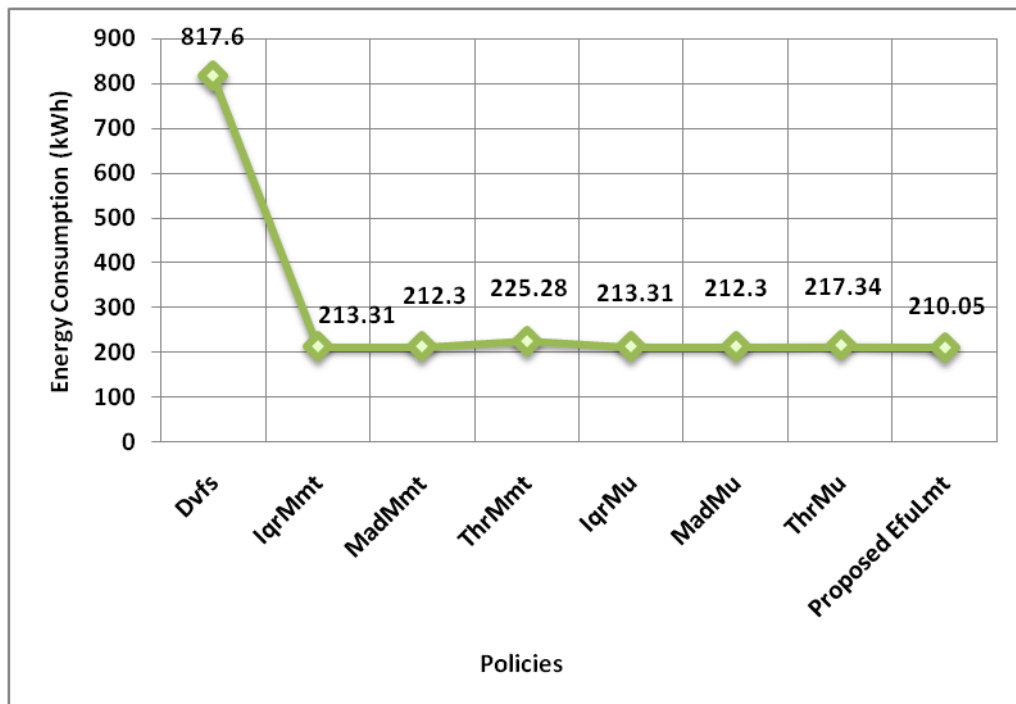


Figure 4: Comparison of Energy Consumption (KWh) of All Policies

Numbers of Host shutdowns are very fewer in case of Dvfs but energy consumption is almost four times the said schemes. The proposed algorithm is being prominently optimizing 2 main parameters and evaluated on basis of these performance metrics. As the overall objective of the proposed framework is to reduce or optimize all the performance parameters defined above. The Average SLA violation in case of DVFS is 0. On the other hand Energy efficient schemes showing very slight Average SLA violation (<10.5%) (Refer Figure 3). Proposed policy showing less SLA violation than three overload detection methods combined with ‘Mmt’ and ‘Mu’. Here the overall reduction in energy consumed by the DCs is showing

significant improvement over few existing policies (see Figure 4). The Average SLA violation is also showing noticeable improvement. The noticeable improvement is due to decline in VM migrations which is also evaluated but not presented as a part of this paper.

## 5. CONCLUSION AND FUTURE WORK

VM consolidation is a rising solution for energy saving in cloud DCs. VM Consolidation is key activity to realize resource allocation in a Cloud DC and power cutback. Also, to meet SLAs while maintaining power-efficient deployment is essential at DCs. The focus of this work is on Quality and

Power aware efficient resource management scheme through dynamic VM consolidation that can be applied in a virtualized DC by a Cloud provider. So, keeping in mind this research mainly focuses on energy-efficient computing and present: (a) survey on various VM consolidation techniques; (b) A proposed novel “Pervasive SLA and Energy Aware VM Consolidation Technique in Cloud DCs”; (c) Performance comparison of existing energy efficient and proposed VM Consolidation Techniques using real workload traces, on basis of various heuristics. The comparison is also done on various parameters with DVFS policy. The proposed policy proved to be reliable as the final values, obtained for maintaining QoS, are optimal.

VM Allocation, Selection, and Placement policies are very important area of research that is not yet explored enough. In future the proposed policies can be tested with varied workloads to make the DCs scalable and reliable in terms of QoS. Also, this work suggests the analysis and modification to energy-efficient policies under dynamic VM consolidation. The ‘Pervasive SLA and Energy aware’ algorithm implementation may prove potential for future workload detection and individual problems of resource consolidation. The proposed policies can be further combined with other energy efficient techniques and tested for performance with real cloud environment, having small number of hosts and VMs.

## 6. REFERENCES

- [1] Wenying Yue and Qiushuang Chen, “Dynamic Placement of Virtual Machines with Both Deterministic and Stochastic Demands for Green Cloud Computing”, Hindawi Publishing Corporation Mathematical Problems in Engineering, Volume 2014, Article ID 613719.
- [2] Belady C. “In the data center, power and cooling costs more than the it equipment it supports” 2007. URL: <http://www.electronics-cooling.com/articles/2007/feb/a3/>.
- [3] <http://www.intel.in/content/dam/www/public/us/en/documents/guides/cloud-computing-virtualization-building-private-iaas-guide.pdf>
- [4] Anton Beloglazov and Rajkumar Buyya, “Adaptive Threshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Data Centers”, MGC '2010, 29 November - 3 December 2010, Bangalore, India. Copyright 2010 ACM 978-1-4503-0453-5/10/11.
- [5] Clark C, Fraser K, Hand S, Hansen JG, Jul E, Limpach C, Pratt I, Warfield A. Live migration of virtual machines. Proceedings of the 2nd Symposium on Networked Systems Design and Implementation (NSDI 2005), USENIX, Boston, MA, USA, 2005.
- [6] Hesham Hassan, Ahmed Shawky Moussa, “Power Aware Computing Survey”, International Journal of Computer Applications (0975 – 8887) Volume 90 – No.3, March 2014.
- [7] N. Bobroff, A. Kochut, and K. Beaty, “Dynamic Placement of Virtual Machines for Managing SLA Violations,” Proc. IFIP/ IEEE 10th Int'l Symp. Integrated Network Management (IM), pp. 119-128, 2007.
- [8] Verma A, Ahuja P, Neogi A. pMapper: Power and migration cost aware application placement in virtualized systems. Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware (Middleware 2008), Springer, Leuven, Belgium, 2008; 243–264.
- [9] Etienne Le Sueur and Gernot Heiser, “Dynamic Voltage and Frequency Scaling: The Laws of Diminishing Returns” url: [https://ts.data61.csiro.au/publications/nicta\\_full\\_text/4158.pdf](https://ts.data61.csiro.au/publications/nicta_full_text/4158.pdf)
- [10] P. Arroba, J. M. Moya, J. L. Ayala, and R. Buyya, “Dynamic voltage and frequency scaling-aware dynamic consolidation of virtual machines for energy efficient cloud data centers,” *Concurrency Comput.: Practice Experience*, vol. 29, 2017, Art. no. e4067.
- [11] Saurabh Kumar Garg , Adel Nadjaran Toosi, Srinivasa K. Gopalaiyengar, Rajkumar Buyya, “SLA-based virtual machine management for heterogeneous workloads in a cloud datacenter”, Elsevier - Journal of Network and Computer Applications 45(2014)108–120.
- [12] Buyya R, Beloglazov A, Abawajy J. Energy-efficient management of data center resources for cloud computing: a vision, architectural elements, and open challenges” in Proceedings of the 2010 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA 2010). Las Vegas, USA, July 2010.
- [13] Anton Beloglazov and Rajkumar Buyya, “Adaptive Threshold-Based Approach for Energy-Efficient Consolidation of Virtual Machines in Cloud Data Centers”, MGC '2010, 29 November - 3 December 2010, Bangalore, India. Copyright 2010 ACM 978-1-4503-0453-5/10/11.
- [14] Anton Beloglazov and Rajkumar Buyya, “Managing Overloaded Hosts for Dynamic Consolidation of Virtual Machines in Cloud Data Centers under Quality of Service Constraints”, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 24, NO. 7, JULY 2013.
- [15] Jung G, Joshi KR, Hiltunen MA, Schlichting RD, Pu C. Generating adaptation policies for multi-tier applications in consolidated server environments. Proceedings of the 5th IEEE International Conference on Autonomic Computing (ICAC 2008), Chicago, IL, USA, 2008; 23–32.
- [16] B. Guenter, N. Jain, and C. Williams, “Managing Cost, Performance, and Reliability Tradeoffs for Energy-Aware Server Provisioning,” Proc. IEEE INFOCOM, pp. 1332-1340, 2011.
- [17] Anton Beloglazov, Rajkumar Buyya, “Optimal online deterministic algorithms and adaptive heuristics for energy and performance efficient dynamic consolidation of virtual machines in cloud data centers”, Wiley InterScience, *Concurr. Comput. : Pract. Exper.*, 24(13):1397-1420, September 2012.
- [18] Anton Beloglazov, Jemal Abawajy, Rajkumar Buyya, “Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing”, *Future Generation Computer Systems* Volume 28, Issue 5, May 2012, pp. 755–768, Elsevier.
- [19] Buyya, R.; Ranjan, R.; Calheiros, R.N., "Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities," *High Performance Computing & Simulation*, 2009.

HPCS '09. International Conference on , vol., no., pp.1,11, 21-24 June 2009.

- [20] Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, Cesar A. F. De Rose, and Rajkumar Buyya, "CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms," *Software: Practice and*

*Experience (SPE)*, Volume 41, Number 1, Pages: 23-50, ISSN: 0038-0644, Wiley Press, New York, USA, January, 2011.

- [21] [http://www.spec.org/power\\_ssj2008/results/res2011q1/power\\_ssj2008-20110124-00338.html](http://www.spec.org/power_ssj2008/results/res2011q1/power_ssj2008-20110124-00338.html)
- [22] [http://www.spec.org/power\\_ssj2008/results/res2011q1/power\\_ssj2008-20110124-00339.html](http://www.spec.org/power_ssj2008/results/res2011q1/power_ssj2008-20110124-00339.html)