



A Survey of Energy Efficient Data Centres in a Cloud Computing Environment

Akshat Dhingra¹, Sanchita Paul²

Department of Computer Science and Engineering, Birla Institute of Technology, Ranchi, India¹

Department of Computer Science and Engineering, Birla Institute of Technology, Ranchi, India²

Abstract: The development of computing systems has always been focused on performance improvements driven by the demand of applications from consumer, scientific and business domains, but the ever increasing energy consumption of computing systems has started to limit further performance growth due to overwhelming energy consumption and carbon dioxide footprints. Hence, the goal of the computer system design has been shifted from performance improvements to power and energy efficiency. Techniques like Dynamic Voltage Frequency Scaling, VM consolidation, Server power switching and resource throttling have been made use of to make cloud computing more power aware and energy efficient [2][3][4][5][6][7]. Dynamic Voltage Frequency Scaling is a technique through which the frequency of the CPU can be dynamically adjusted to reduce energy consumption. VM consolidation is technique whereby VMs residing on different nodes (Servers) are consolidated onto a single node for effective resource utilization and power savings. Server power switching refers to the technique wherein the idle servers could be switched off. Resource throttling could also be done in various ways at the hardware or at the software level in-order to meet the performance requirements and minimize the energy consumption [3]. Algorithms for VM allocation and migration like Greedy, Knapsack, Maximum Bin Packing, Power Expand Min-Max [7] and Minimization Migrations, Highest Potential growth, Random Choice [17] have been used respectively to improve energy efficiency and minimize cost and SLA violations. We therefore review the various techniques mentioned above (and a few others) that have been proposed specifically at the data centre level with the help of virtualization technology. This paper would be concluded with the suggestions on the future research directions.

Keywords: Dynamic Voltage Frequency Scaling, Energy Efficiency, Green Cloud, Resource throttling, Server power switching, Soft scaling, VM Consolidation.

1. INTRODUCTION

The designers have always primarily focused on improving the performance of computing systems and hence the performance has been steadily growing driven by more efficient system design and increasing density of the components described by Moore's law. Although the performance per watt ratio has been constantly rising, the total power draw by computing systems is hardly decreasing. On the contrary, it has been increasing every year that can be illustrated by the estimated average power usage across three classes of servers shown in the table below [19] in Watts/Unit.

Class of Server	2000	2001	2002	2003	2004	2005	2006
Low-end	186	193	200	207	213	219	225
Mid-Range	424	457	491	524	574	625	675
High-end	5534	5832	6130	6428	6973	7651	8163

Table 1: Power Consumption (Watts/Unit)

Apart from the overwhelming operating costs due to high energy consumption, another rising concern is the environmental impact in terms of carbon dioxide (CO₂) emissions caused by this high energy consumption. In 2007, the total carbon footprint of the IT industry including personal computers, mobile phones, and telecom devices and data centres was 830 MtCO₂e, 2% of the estimated total emissions and this figure is expected to grow in the coming years [1]. Therefore, the reduction of power and energy consumption has become a first-order objective in the design of modern computing systems.

There are two possible solutions to make IT Systems greener: 1) improve efficiency or 2) find a plentiful supply of clean, affordable energy. As the latter is still in the realms of science fiction, energy efficiency is where the main focus of research will be in the near future. IT companies are learning that cutting emissions and cutting costs naturally go together, by making systems energy efficient money may be saved automatically. Energy-Aware Computing research is attempting to address this problem. Work in this field is tackling issues ranging from reducing the amount of energy required by a single processor chip to finding the most effective means of cooling a warehouse sized data centre. Cloud Computing



has the potential to have a massive impact, positive or negative, on the future carbon footprint of the IT sector. On the one hand, Cloud Computing datacentres are now consuming 0.5% of all the generated electricity in the world, a figure that will continue to grow as Cloud Computing becomes widespread particularly as these systems are “always-on always-available”. However, the large datacentres required by clouds have the potential to provide the most efficient environments for computing. The growing popularity of cloud computing would therefore drive the cloud providers to build efficient systems in order reduce the total cost of ownership (TOC) and hence improve their green credentials. The main aim of Energy-Aware Computing is to promote awareness of energy consumption at both software and hardware levels and hence consume lesser amount of power. The unique position of Cloud Computing allows this area to be brought into sharper focus and will go some way to improving the carbon footprint of IT now and in the future.

2. ENERGY EFFICIENT CLOUD ARCHITECTURE

Various architectures for energy efficient cloud computing have been proposed. The ones proposed by Buyya et al.[17], Huang et al.[16] and Liu[18] have been shown here for the sake of comparison.

Buyya et al. [17] have proposed a high-level architecture for supporting energy-efficient service allocation in Green Cloud computing infrastructure as shown in the figure below.

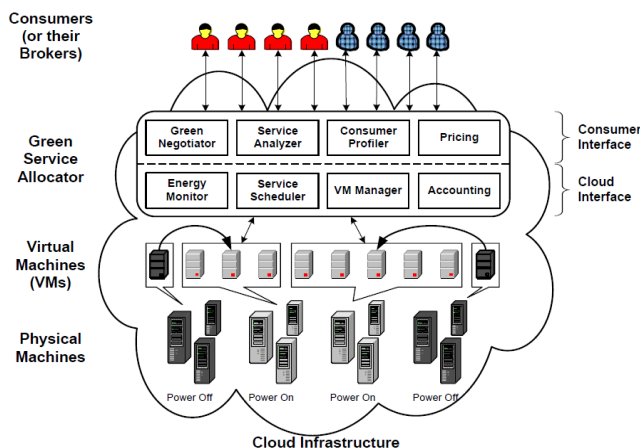


Figure 1: Green Cloud Architecture

The various components involved are:

- Consumers/Brokers:** Cloud consumers or their brokers submit service requests from anywhere in the world to the Cloud
- Green Resource Allocator:** Acts as the interface between the Cloud infrastructure and consumers. It requires the interaction of the following components to support energy-efficient resource management:
 - Green Negotiator: Negotiates with the consumers/brokers to finalize the SLA with specified prices

and penalties (for violations of SLA) between the Cloud provider and consumer.

- Service Analyser: Interprets and analyses the service requirements of a submitted request before deciding whether to accept or reject it.
- Consumer Profiler: Gathers specific characteristics of consumers so that important consumers can be granted special privileges and prioritized over other consumers.
- Pricing: Decides how service requests are charged to manage the supply and demand of computing resources and facilitate in prioritizing service allocations effectively.
- Energy Monitor: Observes and determines which physical machines to power on/off.
- Service Scheduler: Assigns requests to VMs and determines resource entitlements for allocated VMs. It also decides when VMs are to be added or removed to meet demand.
- VM Manager: Keeps track of the availability of VMs and their resource entitlements. It is also in charge of migrating VMs across physical machines.
- Accounting: Maintains the actual usage of resources by requests to compute usage costs.

c) **VMs:** Multiple VMs can be dynamically started and stopped on a single physical machine to meet accepted requests, hence providing maximum flexibility to configure various partitions of resources on the same physical machine to different specific requirements of service requests.

d) **Physical Machines:** The underlying physical computing servers provide hardware infrastructure for creating virtualized resources to meet service demands.

Huang et al. [16] proposed VM-based energy-efficient data centre architecture for cloud computing, as shown in figure below.

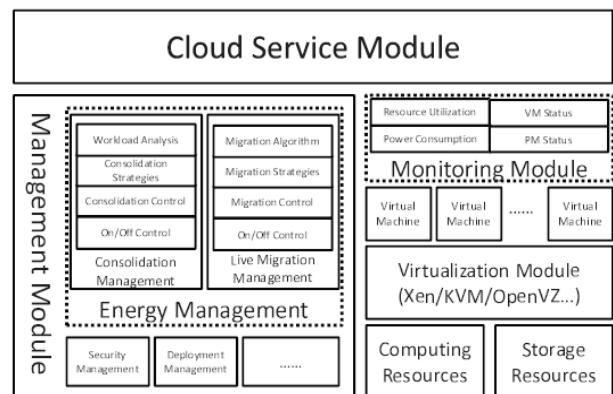


Figure 2: Green Cloud Architecture

The architecture is consisted by four main modules: Virtualization Module, Monitoring Module, Management Module and Cloud Service Module.

- Virtualization module** is an abstraction layer that shields heterogeneous physical resources and provides dynamical, scalable virtual resources to users on demand.
- The Monitoring module** is responsible for monitoring both virtual machines and physical machines; including resource utilization, power consumption and virtual machine status, etc.



c) **Management Module** implements all the management issues in the data centre cloud, including Energy Management Sub module, Security Management Sub module and Deployment Management Sub module, etc.

d) **Cloud Server Module** plays an important role in cloud computing that in which all the resources are provided to the users with services via Internet.

Liu et al. [18] proposed Green cloud architecture as shown below in the figure.

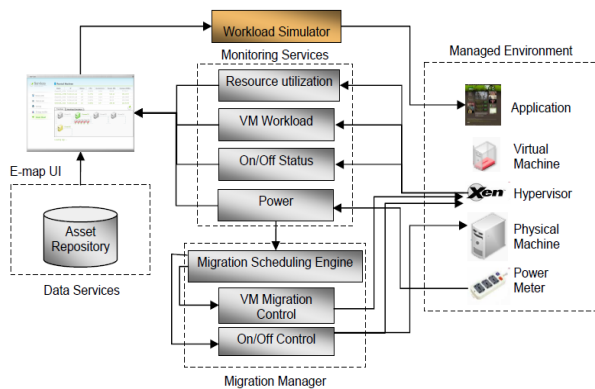


Figure 3: Green Cloud Architecture

- **Monitoring Service** monitors and collects comprehensive factors such as application workload, resource utilization and power consumption, etc.
- **Migration Manager** triggers live migration and makes decision on the placement of virtual machines on physical

servers based on knowledge or information provided by the Monitoring Service

- **Managed Environment** includes virtual machines, physical machines, resources, devices, remote commands on VMs, and applications with adaptive workload, etc.

- **E-Map** is a web-based service with Flash front-end. It provides a user interface (UI) to show the real-time view of present and past system on/off status, resource consumption, workload status, and temperature and energy consumption in the system

- **Workload Simulator** accepts user instructions to adapt workload, e.g. CPU utilization, on servers, and enables the control of Migration Manager under various workloads.

- **Asset Repository** is a database to store the static server information, such as IP address, type, CPU configuration, memory setting, and topology of the servers.

- **The Green Cloud IDC** can view up-to-date status of resources, configure their applications, allocate resources, and experience the live management system.

From the various proposed architectures explained above, there is one broad fact that must be observed, that is, there is an abstraction layer that shields all the physical resources and provides dynamical and scalable resources to the users on demand. This layer has several components that together participate for the purpose of minimizing the cost and hence energy and carbon footprint. A brief summary of the above mentioned architectures has been given below in a tabular form [Table 2].

Table 2: Comparative study of the various architectures for energy efficient data centres

Architectures with author name and reference number =>	Buyya et al. [17]	Liu et al. [15]	Huang et al. [16]
Entities Involved	Consumers, Green Resource Allocator, Virtual machines, Physical machines	Data services, Migration manager, Monitoring Services, Managed Environment	Cloud Service module, Management module, Virtualization module, Monitoring Module
Components involved to improve energy efficiency	Green Negotiator, Service Analyzer, Consumer profiler, Pricing, Energy Monitor, Service Scheduler, VM Manager, Accounting	Migration Scheduling Engine, VM Migration Control, ON/OFF control, Resource utilization, VM workload, Workload Simulator, Asset Repository	Workload Analysis, ON/OFF Control, Migration Strategies, Resource Utilization, Power Consumption

3. VIRTUALIZATION TECHNOLOGY AND THE MINIMIZATION OF ENERGY CONSUMPTION

Virtualization may refer to many things, but in this context we will take it to mean hardware virtualization allowing multiple guest operating systems to run on a single node. Hardware virtualization hides the underlying

computing system to present an abstract computing platform by using a hypervisor. In data centres, the number of physical machines can be reduced using virtualization by consolidating VMs onto shared servers which helps improve the efficiency of IT systems. The advantages are simple, it allows multiple virtual machines to be run on a single physical machine in order to provide



more capability and increase the utilization level of the hardware. In-order to make the most out of Virtualization to save energy, the Cloud computing environment should comply with the following requirements:

- Virtualization of the infrastructure to support hardware and software heterogeneity, optimum resource utilization and simplify the resource provisioning.
- Application of VM migration to continuously adapt the allocation and quickly respond to changes in the workload.
- Ability to handle multiple applications with different SLA requirements owned by multiple users.
- Guaranteed meeting of the QoS requirements for each application.
- Support for different kind of applications, mixed workloads.
- Decentralization and high performance of the optimization algorithm to provide scalability and fault tolerance.
- Optimization considering multiple system resources, such as CPU, memory, disk storage and network interface.

- Intelligent consolidation and live migration mechanism to improve the energy efficiency.
- Designing algorithms to reduce not only the power consumption but also the heat dissipation.

4. EFFECTIVE RESOURCE UTILIZATION TO MINIMIZE ENERGY CONSUMPTION WITH THE HELP OF VIRTUALIZATION

Virtualization technology allows one to create several VMs on a physical server thereby reducing the amount of hardware in use. Various techniques for effective resource utilization have been proposed and some of which have been summarized in a tabular form below [Table 3].

Table 3: Various techniques proposed to make data centres more energy efficient

Paper title with reference No.	Proposed techniques	Virtualization	Goal	Limitations
Autonomic resource management in virtualized data centers using fuzzy logic based approaches. [8]	Local and global controllers making use of fuzzy logic-based approaches to minimize the resource utilization	Yes	Maximize resource utilization, profit and minimize cost.	Does not consider the cost of migrating virtual machines and hence not applicable in scale out datacenters.
VirtualPower: Coordinated Power Management in Virtualized Enterprise Systems. [2]	DVFS, soft scaling, VM consolidation, server power switching, local and global level policies.	Yes	Minimize energy consumption, satisfy performance requirements	Does not include mechanisms for VM-specific power throttling and power allocation, lacks a detailed description of the global policies used.
Coordinated Multi-level Power Management for the Data Center [3]	DVFS, VM consolidation, server power switching	Yes	Minimize power consumption, minimize performance loss, while meeting power budget	Less focus in performance and cooling domains.
Power and Performance Management of Virtualized Computing Environments via Lookahead Control [4]	DVFS, VM consolidation, server power switching	Yes	Minimize power consumption, minimize performance loss	Processing rate of VMs with different CPU Share must be known in advance. Due to complexity of the model, the optimization controller execution time is high.



Resource Allocation using Virtual Clusters.[5]	Resource throttling	Yes	Maximize resource utilization, satisfy performance requirements	Resource needs should be known in advance. Assumes that VM instances are CPU bound but in reality they may have composite needs.
Multi-Tiered On-Demand Resource Scheduling for VM-Based Data Center [6]	Resource throttling	Yes	Maximize resource utilization, satisfy performance requirements	Optimization of only CPU and no other resource. Resource needs should be known in advance.
Shares and Utilities based Power Consolidation in Virtualized Server Environments. [7]	DFVS, soft scaling	Yes	Minimize power consumption, minimize performance loss	VM migration not applied and allocation is static, requires the knowledge of application priorities in advance.
pMapper: Power and Migration Cost Aware Application Placement in Virtualized Systems [9]	DVFS, VM consolidation, server power switching	Yes	Minimize power consumption, minimize performance loss	No consideration of memory, network bandwidth and advanced application of idle states.
Resource pool management: Reactive versus proactive [10]	VM consolidation, server power switching	Yes	Maximize resource utilization, satisfy performance requirements	Reactive migration controller or proactive workload placement controller is not adequate for effective resource pool management.
GreenCloud: Energy-Efficient and SLA-based Management of Cloud Resources.[11]	Leveraging heterogeneity of Cloud data centers, DVFS, Green resource allocator.	Yes	Minimize energy consumption, satisfy performance requirements	Aggressive VM consolidation of VMs may lead to performance degradation and SLA violations.
Energy Efficient Resource Management in Virtualized Cloud Data Centers. [15]	Optimization over multiple system resources, network optimization and thermal optimization. Local and Global managers for VM placement and continuous optimizations.	Yes	Minimize power consumption and handle strict QoS requirements	Proposal and hence lacks implementation on a real system and experimental evaluation.
Virtual Machine Based Energy-Efficient Data Center Architecture for Cloud Computing: A Performance Perspective. [16]	Server consolidation, live migration.	Yes	Minimize energy consumption through efficient server consolidation and live VM migration.	Experimental and lacks a concrete solution to minimize energy consumption.



Energy Efficient Allocation of Virtual Machines in Cloud Data Centers. [14]	Effective dynamic relocation of VMs, Minimization Migration, Highest Potential Growth, Random Choice, DVFS	Yes	Minimize power consumption, satisfy performance requirements	Considers CPU only as a resource and not disk storage, network bandwidth as they also consume significant amount of power.
---	--	-----	--	--

6. CONCLUSIONS AND FUTURE DIRECTIONS

With the growing demand of cloud computing, the energy efficiency has become one of the most important design requirements, as they continue to draw enormous amounts of electrical power.

At present the IT infrastructures contribute to about 2% of the total Carbon Dioxide footprints. Therefore, there is an urgent need for efficient techniques and algorithms to manage the computing resources, otherwise the energy consumption and carbon dioxide emissions are expected to rapidly grow. We have reviewed the various power management schemes at the data center level with the

help of virtualization. Most of the work here aims to optimize no other system resource except for CPU when there are other resources like memory and network bandwidth that also consume power and hence demand consideration. Also transition time for switching power states of the resource and VM migration overhead are not handled leading to performance degradation. The future directions in each of the papers that have been mentioned above in the table have been summarized below in Table 4 along with the resource(s) that they are trying to optimize.

Table 4: Summary of the future work considering the resource under optimization

Paper Title with Reference No.	System Resources	Future Work
Autonomic resource management in virtualized data centers using fuzzy logic based approaches. [8]	CPU	Consider the cost of VM migration in the proposed profit model.
VirtualPower: Coordinated Power Management in Virtualized Enterprise Systems. [2]	CPU	Include abstractions and mechanism support for VM-specific power throttling and power allocation.
Coordinated Multi-level Power Management for the Data Center. [3]	CPU	Include coordination with the equivalent spectrum of solutions in the <i>performance</i> and <i>cooling</i> domains.
Power and Performance Management of Virtualized Computing Environments via Lookahead Control. [4]	CPU	Find solutions to decrease the optimization controller execution time. Optimization through fuzzy logic.
Resource Allocation using Virtual Clusters. [5]	CPU	Account for network and memory usage. Make use of techniques like fuzzy logic to predict the resource needs.
Multi-Tiered On-Demand Resource Scheduling for VM-Based Data Center. [6]	CPU, Memory	Analysis of the potential and overhead caused by each tier of the multi-tiered resource scheduling.
Shares and Utilities based Power Consolidation in Virtualized Server Environments. [7]	CPU	Account for network and memory requirements in its power consolidation algorithms.



pMapper: Power and Migration Cost Aware Application Placement in Virtualized Systems. [9]	CPU	Advanced application of idle states and consideration of network bandwidth and memory as resource.
Resource pool management: Reactive versus proactive. [10]	CPU, Memory	Evaluation of other instances of controllers and management policies. Develop management policies that react well to more kinds of workloads and different kinds of simulated failures.
Green Cloud: Energy-Efficient and SLA-based Management of Cloud Resources. [11]	CPU	Optimization of virtual network topologies. QoS based resource selection and provisioning.
Energy Efficient Resource Management in Virtualized Cloud Data Centers. [15]	CPU	Fine grained network and thermal optimizations.
Virtual Machine Based Energy-Efficient Data Center Architecture for Cloud Computing: A Performance Perspective [16].	CPU	Include design and implementation of intelligent consolidation and live migration mechanism to improve the energy efficiency.
Energy Efficient Allocation of Virtual Machines in Cloud Data Centers. [14]	CPU	Consideration of multiple system resource in reallocation decisions, such as network interface and disk storage. Setting the utilization thresholds dynamically according to a current set of VMs allocated to a host. Decentralization of the optimization algorithms to improve scalability and fault tolerance.

For the future research work we suggest to develop intelligent techniques to manage the network resources efficiently. One of the ways to achieve this for virtualized data centers is to continuously optimize network topologies established between VMs, and thus reduce network communication overhead and load of network devices.

Also, efficient workload distribution across geographically distributed data centers can enable the reallocation of the workload to a place where renewable sources of energy for cooling and operation are available (e.g. solar energy during daytime across different time zones, efficient cooling due to climate conditions). However, it may also be noted that the renewable sources might not always be available making the solution susceptible to failures, so in such cases the workload could be re-directed to the places where the energy is comparatively costly and maybe non-renewable but always available. This is done to meet the performance requirements and make the system energy efficient. Other important directions are, providing fine grained user's control over power consumption / CO2 emissions in Cloud environments and support for flexible SLA negotiated between resource providers and users.

It would also be interesting to access the potential of software optimizations in terms of energy consumption. Currently software engineers optimize for traditional performance, but guidelines for energy could be drafted to aid the long-term efficiency of computing systems.

REFERENCES

- [1] E. Farnworth and J.C. Castilla--rubio, "SMART 2020: Enabling the low carbon economy in the information age," Group.
- [2] R. Nathuji and K. Schwan, —Virtualpower: Coordinated power management in virtualized enterprise systems, ACM SIGOPS Operating Systems Review, vol. 41, no. 6, pp. 265–278, 2007.
- [3] R. Raghavendra, P. Ranganathan, V. Talwar, Z. Wang, and X. Zhu, —No "power" struggles: Coordinated multi-level power management for the data center, SIGARCH Computer Architecture News, vol. 36, no. 1, pp. 48–59, 2008.
- [4] D. Kusic, J. O. Kephart, J. E. Hanson, N. Kandasamy, and G. Jiang, —Power and performance management of virtualized computing environments via lookahead control, Cluster Computing, vol. 12, no. 1, pp. 1–15, 2009.
- [5] M. Stillwell, D. Schanzenbach, F. Vivien, and H. Casanova, —Resource allocation using virtual clusters,| in Proceedings of the 9th IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid 2009), Shanghai, China, 2009, pp. 260–267.
- [6] Y. Song, H. Wang, Y. Li, B. Feng, and Y. Sun, —Multi-Tiered On-Demand resource scheduling for VM-Based data center,| in Proceedings of the 9th IEEE/ACM International Symposium on Cluster Computing and the Grid (CCGrid 2009), Shanghai, China, 2009, pp. 148–155.



- [7] M. Cardosa, M. Korupolu, and A. Singh, —Shares and utilities based power consolidation in virtualized server environments, | in Proceedings of the 11th IFIP/IEEE Integrated Network Management (IM 2009), Long Island, NY, USA, 2009.
- [8] Autonomic resource management in virtualised data centers using fuzzy logic based approaches Jing Xu, Ming Zhao, José Fortes, Robert Carpenter, Mazin Yousif in Journal Cluster Computing archive Volume 11 Issue 3, September 2008 Pages 213 – 227.
- [9] A. Verma, P. Ahuja, and A. Neogi, —pMapper: power and migration cost aware application placement in virtualized systems, in Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware. Springer-Verlag New York, Inc., 2008, pp. 243–264.
- [10] D. Gmach, J. Rolia, L. Cherkasova, and A. Kemper, —Resource pool management: Reactive versus proactive or let's be friends, Computer Networks, 2009.
- [11] R. Buyya, A. Beloglazov, and J. Abawajy, —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 12-15, 2010.
- [12] D. Greenwood, A. Khajeh---hosseini, J. Smith, and I. Sommerville, "The Cloud Adoption To2lkit: Addressing the Challenges of Cloud Adoption in Enterprise," Challenges, 2010.
- [13] X. Fan, W. D. Weber, and L. A. Barroso, —Power provisioning for a warehouse-sized computer,| in Proceedings of the 34th Annual International Symposium on Computer Architecture (ISCA 2007). ACM New York, NY, USA, 2007, pp. 13–23.
- [14] Anton Beloglazov and Rajkumar Buyya - Energy Efficient Allocation of Virtual Machines in Cloud Data Centers in 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 2010.
- [15] Anton Beloglazov and Rajkumar Buyya- Energy Efficient Resource Management in Virtualized Cloud Data Centers in 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 2010.
- [16] Kejiang Ye, Dawei Huang, Xiaohong Jiang, Huajun Chen, Shuang Wu- Virtual Machine Based Energy-Efficient Data Center Architecture for Cloud Computing: A Performance Perspective in IEEE/ACM International Conference on Green Computing and Communications, 2010 & IEEE/ACM International Conference on Cyber, Physical and Social Computing, 2010.
- [17] Rajkumar Buyya, Anton Beloglazov, and Jemal Abawajy- Energy-Efficient Management of Data Center Resources for Cloud Computing: A Vision, Architectural Elements, and Open Challenges Proceedings of the 2010 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA 2010), Las Vegas, USA, July 12-15, 2010.
- [18] Liang Liu, Hao Wang, Xue Liu, Xing Jin, WenBo He, QingBo Wang, Ying Chen1-GreenCloud: A New Architecture for Green Data Center in ICAC-INDST '09 Proceedings of the 6th international conference industry session on Autonomic computing and communications industry session, Pages 29-38.
- [19] J. G. Koomey, —Estimating total power consumption by servers in the US and the world, Oakland, CA: Analytics Press. February 15, 2007.