978-81-931039-6-8

DESIGNING THE ARCHITECTURE OF INTELLIGENT AGENT BASED FRAMEWORK IN IMPROVING CLOUD RESOURCE MANAGEMENT (CRM) POLICIES

Preety[1], Dr. Surjeet Dalal[2], Deepika[3]

[1] *Phd Scholar from Dr. K.N. Modi University (India)*

[2]*Associate Professor, CSE Department, SRM University, Haryana (India)*

[3]*Teacher in Modern School Delhi (India)*

ABSTRACT

Cloud Computing is a technology through which IT resources can be used via Internet from any location with pay-as you-go model without spending a large amount of capital on purchasing and installing them. An Intelligent Agent is a standalone/autonomous entity (program) that can observe and react flexibly in a dynamic and uncertain environment for achieving some goal in place of human being. In a Multi agent system multiple agents interact with each other for solving a particular problem. For successful interaction of agents, they require to cooperate, coordinate and negotiate with each other. This paper discusses Cloud Computing, its uses through a service-oriented interface to offer on demand services, gives overview of Intelligent Agents in Cloud Computing, the role of agents in Cloud Computing and how this Intelligent Agent can help in facilitating better services to Cloud Computing. We have also discussed the design of the system and implementation of an application in Cloud System using Intelligent Agents.

I. INTRODUCTION

In computer science, the Cloud Computing describes a type of outsourcing of computer services which are similar to the way in which the supply of electricity is out sourced. The idea behind cloud computing is similar that the user can simply use storage computing power or specially crafted development environments without having to worry how these work internally. Cloud computing is a system architecture model for internet based computing. It is the development and use of computer technology on the internet. Data centers and cloud computing services providers hope that the extensive adoption of the cloud will bring them more revenue and they are dynamically promoting the technology.

Cloud computing is a type of computing that relies on sharing computing resources rather than having local servers or personal devices to handle applications. The ‘cloud’ in cloud computing can be defined as the set of hardware, network, storage services and interface that combine to deliver aspects of computing as a service.

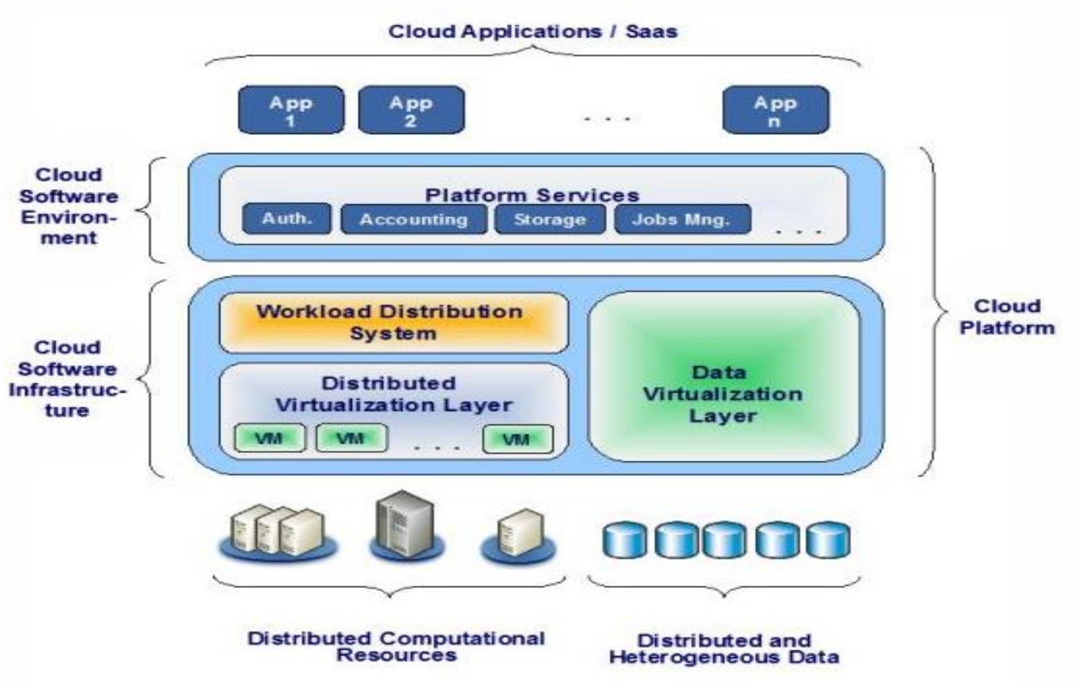

Figure 1. Cloud Architecture

II. SERVICE MODELS

2.1 Infrastructure as a Service (IaaS)

In the most basic cloud-service model, providers of IaaS offer computers - physical or virtual machines - and other resources. A hypervisor, such as KVM, runs the virtual machines as guests. Pools of hypervisors within the cloud operational support-system can support large numbers of virtual machines and the ability to scale services up and down according to customers varying requirements.

2.2 Platform as a Service (PaaS)

In the PaaS model, cloud providers deliver a computing platform typically including operating system, programming language execution environment, database, and web server. Application developers can develop and run their software solutions on a cloud platform without the cost and complexity of buying and managing the underlying hardware and software layers.

2.3 Software as a Service (SaaS)

In the business model using software as a service (SaaS), users are provided access to application software and databases. Cloud providers manage the infrastructure and platforms that run the applications. SaaS is sometimes referred to as "on-demand software" and is usually priced on a pay-per-use basis. SaaS providers generally price applications using a subscription fee.

III. DEPLOYMENT MODELS

a. Private cloud

Undertaking a private cloud project requires a significant level and degree of engagement to virtualize the business environment, and requires the organization to reevaluate decisions about existing resources. When done right, it can improve business, but every step in the project raises security issues that must be addressed to prevent serious vulnerabilities.

b. Public cloud

A cloud is called a 'Public cloud' when the services are rendered over a network that is open for public use. Technically there is no difference between public and private cloud architecture, however, security consideration may be substantially different for services like applications, storage, and other resources that are made available by a service provider for a public audience and when communication is effected over a non-trusted network.

c. Hybrid cloud

Hybrid cloud is a composition of two or more clouds like private, community or public that remain unique entities but are bound together, offering the benefits of multiple deployment models. Such composition expands deployment options for cloud services, allowing IT organizations to use public cloud computing resources to meet temporary needs.

IV. ISSUES IN CLOUD COMPUTING

Energy Efficiency: Cloud computing is a vast emerging IT field that is in rapid form of expansion but cloud computing energy consumption is one of the area that needs to be focused.

Security: Although the cloud computing service provider will provide with security essentials likes data storage and transmission encryption, authentication and authorizations but vulnerability of remote data access, Session ridding, virtual machine escape, Data storage, Criminal hackers, thieves and corrupt employees causes a lot of concern.

Reliability: Some people worry also about whether a cloud service provider is financially stable and whether their data storage system is trustworthy.

Ownership: Once data has been relegated to the cloud, some people worry that they could lose some or all of their rights or be unable to protect the rights of their customers.

Data Backup: Cloud providers employ redundant servers and routine data backup processes, but some people worry about being able to control their own backups.

Data Portability and Conversion: As service competition grows and open standards become established, the data portability issue will ease, and conversion processes will become available supporting the more popular cloud providers. Worst case, a cloud subscriber will have to pay for some custom data conversion.

Multiplatform Support: More an issue for IT departments using managed services is how the cloud-based service integrates across different platforms and operating systems. Usually, some customized adaption of the service takes care of any problem.

Intellectual Property: A company invents something new and it uses cloud services as part of the invention. Is the invention still patentable? Or there can be issues like cloud service provider can make claim for that invention or leak the information to the competitor.

V. REVIEW OF LITERATURE

There are energy cost concerns in every cloud service providers, as data centers are the heavy duty power thirsty equipments. The data center has changed significantly as the development of information technology which has enabled it to become the critical nerve center of today's enterprise. As business difficulty increase, so does the number of data center facilities which house a rising amount of powerful IT equipment. Data center managers around the world are running into resource limits related to power, cooling, and space, building the resource efficiency of data centers an important topic of debate. As a global consortium comprised of end-users, policy makers , technology providers , facility architects, and utility companies. The services can be of any type e.g. Infrastructure as a Service (IaaS) e.g. Platform as a Service (PaaS) and Software as a Service (SaaS) The major benefits of cloud data centers includes the tradition of financial system of scale to pay back the cost of ownership and the cost of system maintenance over a large number of machines. Customers will be able to access infrastructure and data from a cloud anywhere from the world. With the rapid growth of cloud data centers, the energy consumed by data centers is huge and straightforwardly associated to the number of hosted servers and their workload. It has extremely increased over the past ten years. The power consumption of data centers has huge impacts on the environment. The amount of electricity consumed by data centers worldwide dramatically grew also the electricity cost mostly in developing countries has already under a hike.

Moreno-Vozmediano et al. (2013) stated that Cloud computing played a major role in the future Internet of Services, enabling on-demand provisioning of applications, platforms, and computing infrastructures. The cloud community addressed several technology challenges to turn that vision into reality. They specified the issues relate to deploying future infrastructure-as-a-service clouds and include efficiently managing such clouds to deliver scalable and elastic service platforms on demand, developing cloud aggregation architectures and technologies that let cloud providers collaborated and interoperated, and improving cloud infrastructures' security, reliability, and energy efficiency.

Younge et al. (2011) presented new framework that provided efficient green enhancements within a scalable Cloud computing architecture. Using power-aware scheduling techniques, variable resource management, live migration, and a minimal virtual machine design, overall system efficiency vastly improved in a data center based Cloud with minimal performance overhead The notion of Cloud computing had not only reshaped the field of distributed systems but also fundamentally changed how businesses utilize computing these day. The Cloud computing provided many advanced features which has some shortcomings such as the relatively high operating cost for both public and private Clouds.

Agrawal et al. (2012) presented the cloud computing in which resources were available on the temporary basis or in the leased. The resources were in the form or combination of Software and Hardware the customer utilized this resources. Problem area was when there were more users for single cloud resources at instance of time then how they could be synchronized the resources scheduling for more than one user. However, managing several resources, potentially with different architectures, was difficult for users. Another difficulty was optimally scheduling applications in such environment. In this paper they were giving the strategy how the resource managed in cloud environment. Here Resource management meant single cloud having cluster of functional server, so that they could schedule the cloud resources for number of different user. Then they compared the cloud computing with grid computing, Cloud computing evolved from grid computing and provided on-demand resource provisioning. Grid computing might or might not be in the cloud depending on what type of users are using it. If the users were systems administrators and integrators, users cared how things are maintained in the cloud.

Beloglazov et al. (2010) proposed an energy efficient resource management system for virtualized Cloud data centers that reduced operational costs and provided required Quality of Service (QoS). Energy savings were achieved by continuous consolidation of VMs according to current utilization of resources, virtual network topologies established between VMs and thermal state of computing nodes. They presented first results of simulation-driven evaluation of heuristics for dynamic reallocation of VMs using live migration according to current requirements for CPU performance. The results showed that the proposed technique brings substantial energy savings, while ensuring reliable QoS. This paper justified further investigation and development of the proposed resource management system.

Buyya et al. (2011) presented vision, challenges, and architectural elements of SLA-oriented resource management. The proposed architecture supported integration of market based provisioning policies and virtualisation technologies for flexible allocation of resources to applications. The performance results obtained from their working prototype system showed the feasibility and effectiveness of SLA-based resource provisioning in Clouds. Cloud computing systems promised to offer subscription-oriented, enterprise-quality computing services to users worldwide. With the increased demand for delivering services to a large number of users, they needed to offer differentiated services to users and meet their quality expectations. Existing resource management systems in data centers were yet to support Service Level Agreement (SLA)-oriented resource allocation, and thus needed to be enhanced to realize cloud computing and utility computing. In addition, no work had been done to collectively incorporate customer-driven service management, computational risk management, and autonomic resource management into a market-based resource management system to target the rapidly changing enterprise requirements of Cloud computing.

Wuhib et al. (2012) addressed the problem of dynamic resource management for a large-scale cloud environment. Their contribution included outlining the distributed middleware architecture and presented one of its key elements: a gossip protocol that ensured fair resource allocation among sites or applications, dynamically adapted the allocation to load changes and scaled both in the number of physical machines and sites or applications. They formalized the resource allocation problem as that of dynamically maximizing the cloud utility under CPU and memory constraints. They first presented a protocol that computed an optimal solution without considering memory constraints and proved correctness and convergence properties. Then, they extended that protocol to provide an efficient heuristic solution for the complete problem, which includes minimizing the cost for adapting an allocation. The protocol continuously executed on dynamic, local input and did not require global synchronization, as other proposed gossip protocols do. They evaluated the heuristic protocol through simulation and found its performance to be well-aligned with their design goals.

Lee et al. (2013) integrated the system resource, utilized the resource flexibly, saved the energy consumption in the cloud computing and meted the requirements of users in the cloud computing environment, one of the positive solutions was to apply the virtualization technology. A dynamic resource management with energy saving mechanism had been proposed in this paper which presented a method of dynamic voltage scaling for dynamic adjustment of resources by inspecting CPU utilization in the cloud computing environment. The voltage of the idle or light loaded computer could be reduced and heavy loaded works can be migrated to those machines with lighter loading for achieving the purpose of energy saving. The experimental results showed that at the user unaware situation, energy consumption could be saved significantly by applying the proposed mechanism.

Gulati et al. (2011) shed light on some of the key issues in building cloud-scale resource management systems, based on five years of research and shipping cluster resource management products. They discussed various techniques to provide large scale resource management, along with the pros and cons of each technique. They hoped to motivate future research in this area to develop practical solutions to these issues. Managing resources at large scale while providing performance isolation and efficient use of underlying hardware was a key challenge for any cloud management software. Most virtual machine (VM) resource management systems like VMware DRS clusters, Microsoft PRO and Eucalyptus, did not currently scale to the number of hosts and VMs supported by cloud service providers. In addition to scale, other challenges include heterogeneity of systems, compatibility constraints between virtual machines and underlying hardware, islands of resources created due to storage and network connectivity and limited scale of storage resources.

Vinothina et al. (2012) stated that Cloud computing had become a new age technology that had got huge potentials in enterprises and markets. Clouds could make it possible to access applications and associated data from anywhere. Companies were able to rent resources from cloud for storage and other computational purposes so that their infrastructure cost could be reduced significantly. Further they could make use of company-wide access to applications, based on pay-as-you-go model. Hence there was no need for getting licenses for individual products. However one of the major pitfalls in cloud computing was related to optimizing the resources being allocated. Because of the uniqueness of the model, resource allocation was performed with the objective of minimizing the costs associated with it. The other challenges of resource allocation were meeting customer demands and application requirements. In this paper, various resource allocation strategies and their challenges were discussed in detail. It was believed that this paper would benefit both cloud users and researchers in overcoming the challenges faced.

Jhawar et al. (2012) addressed the problem of guaranteeing security, with additional consideration on reliability and availability issues, in the management of resources in Cloud environments. They investigated and formulated different requirements that users or service providers might wish to specify. Their framework allowed providers to impose restrictions on the allocations to be made to their hosts and users to express constraints on the placement of their virtual machines. User’s placement constraints might impose restrictions in performing allocation to specific locations, within certain boundaries, or depending on some conditions e.g., requiring a VM to be allocated to a different host with respect to other VMs). Their approach for VM allocation went beyond the classical like performance or cost-oriented resource consumption to incorporate the security requirements specified by users and providers.

VI. INTELLIGENT AGENT

An intelligent agent is a software that assists people and act on their behalf. Intelligent agents work by allowing people to delegate work that they could have done, to the agent software. Agents can perform repetitive tasks, remember things you forgot, intelligently summarize complex data, learn from you and even make recommendations to you.

All agents are autonomous, which means that an agent has control over its own actions. All agents are also goal-driven. Agents have a purpose and act accordance with that purpose. There are several ways of making goals known to an agent, and are listed below:

- An agent could be driven by a script with pre-defines action which would then define the agent's goals.
- An agent could also be a program and as long as the program is driven by goals and has other characteristics of agents.
- An agent could also be driven by rules, and the rules would define the agent's goals.
- There is also embedded agent goals, such as "planning" methodologies, and in some cases the agent could change its own goals over time.

An agent could also senses changes in its environment and responds to these changes. This characteristic of the agent is at the core of delegation and automation. For example, you tell your assistant "when x happens, do y" and the agent is always waiting for x to happen. An agent continue to work even when the user is gone, which means that an agent could run on a server, but in some cases, an agent run on the user systems.

In a Multi-Agent System, agents are social, this means that they communicate with other agents. Some agents learn or change their behavior base on their previous experiences. Some agents are mobile, meaning they move from machine to machine to be closer to data they may need to process and do so without network delays. Finally, some agents attempt to be believable, such that they are represented as an entity visible or audible to the user and may even have aspects of emotion or personality

Some of the agent communication languages include KQML (Knowledge Query and Manipulation Languages), AOP (Agent Oriented Programming) and Agent Talk.

- KQML is a language and protocol used for exchanging information and knowledge. KQML is both a message format and a message-handling protocol to support run-time knowledge sharing among agents. KQML can be used as a language for an application program to interact with an intelligent system or for two or more intelligent systems to share knowledge in support of cooperative problem solving.
- AOP is an interpreter for programs written in a language called AO. AO is a programming language for the paradigm of Agent-Oriented Programming. It is currently under development at Stanford.
- Agent Talk is a coordination protocol description language for multi agent systems. Agent Talk allows coordination protocols to be defined incrementally and to be easily customized to suit application domains by incorporating an inheritance mechanism.

VII. ARCHITECTURE OF INTELLIGENT AGENT BASED FRAMEWORK

We provide a brief summary of literature survey indicates that with traditional resource management in cloud computing system. Since the fluctuating electricity cost is very high for different data centers locating in different geographical positions. Therefore an algorithmic efficiency must be taken into account with other parameters.

Figure 2. Cloud Agent

To develop a system that will be cost based virtual machine migration.

- To establish an efficient resource allocation System for minimizing Service Level Violations (SLAs).
- To develop a QoS system for users in terms of cost efficiency and response time i.e. time taken from the Cloud to respond users request..

We review previous offerings to the discussion of resource management which provide efficiency of cloud computing, provide a working definition of cloud computing and discuss its importance, which will grow as the technology matures and becomes well known. Further, the main contribution of this work is to minimize the energy cost of a cloud which has huge impact on the economy of the Cloud Service Providers. Therefore this work focuses

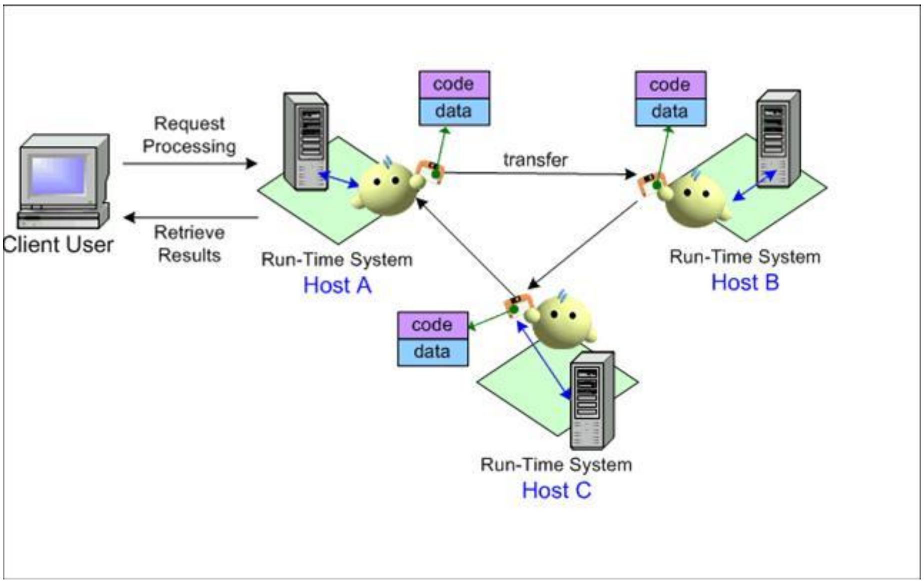

Figure 3. Cloud Environment

VIII.CONCLUSION

The association of Agent and the Cloud is beneficial for both the parties: Cloud User and Cloud Providers. The proposal suggests an AGENT for cloud computing services user. In today's scenario there might be very few service providers for cloud computing and there might be very few services provided by them. But what if the service providers and the services will increase in future? It would be difficult to find the optimized services on our own. So we require some kind of automated procedure that will take care of the things no matter how many service providers are there and how many services they are providing.

REFERENCES

[1]. Moreno-Vozmediano, R.; Montero, R.S.; Llorente, I.M., "Key Challenges in Cloud Computing: Enabling the Future Internet of Services," *Internet Computing, IEEE* , vol.17, no.4, pp.18,25, July-Aug. 2013.

[2]. Andrew J. Younge, Gregor von Laszewski, Lizhe Wang, "Efficient Resource Management for Cloud Computing Environments", Proceeding GREENCOMP '10 Proceedings of the International Conference on Green Computing, Pages 357-364, IEEE Computer Society Washington, DC, USA ©2010.

[3]. Rajkumar Buyya, Rajiv Ranjan, "Federated resource management in grid and cloud computing systems", Future Generation Computer Systems 26 (2010) 1189_1191.

[4]. Preeti Agrawal, Yogesh Rathore, "Resource Management In Cloud Computing With Increasing Dataset", International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 6, June 2012.

[5]. Anton Beloglazov and Rajkumar Buyya, "Energy Efficient Resource Management in Virtualized Cloud Data Centers", Proceeding 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 826-834

[6]. Rajkumar Buyya, Saurabh Kumar Garg1, and Rodrigo N. Calheiros, "SLA-Oriented Resource Provisioning for Cloud Computing: Challenges, Architecture, and Solutions", 2011 IEEE International Conference on Cloud and Service Computing, 1-10

[7]. Fetahi Wuhib and Rolf Stadler, "A Gossip Protocol for Dynamic Resource Management in Large Cloud Environments", IEEE transactions on network and service management (TNSM), Vol. 9, No. 2, June 2012, 1-10.

[8]. Ajay Gulati, "Cloud-Scale Resource Management: Challenges and Techniques", Workshop on Hot Topics in Management of Internet, Cloud, and Enterprise Networks and Services, USA, 2011

[9]. V.Vinothina, "A Survey on Resource Allocation Strategies in Cloud Computing", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 3, No.6, 2012.

[10]. Jhawar, R.; Piuri, V.; Samarati, P., "Supporting Security Requirements for Resource Management in Cloud Computing," Computational Science and Engineering (CSE), 2012 IEEE 15th International Conference on , vol., no., pp.170,177, 5-7 Dec. 2012.

[11]. Alexandre Kandalintsev, "Advanced Cloud Resource Management For Performance and Power Saving", CHANGE & OFELIA Summer school, November 7-11. Berlin, Germany.

[12]. Andreas Berl, Erol Gelenbe, Marco di Girolamo, Giovanni Giuliani, Hermann de Meer, Minh Quan Dang and Kostas Pentikousis ," Energy-Efficient Cloud Computing" in Advance Access publication on August 19, 2009, Published by Oxford University Press on behalf of The British Computer Society.

[13]. Anton Beloglazov, Rajkumar Buyya, Young Choon Lee, and Albert Zomaya " A Taxonomy and Survey of Energy-Efficient Data Centers and Cloud Computing Systems" in Proc. of the ACM/IEEE Conf. on Supercomputing (SC), Seattle, WA, 2005, pp. 1–9.

[14]. Anton Beloglazov, Jemal Abawajy, Rajkumar Buyya ," Energy-aware resource allocation heuristics for efficient management of data centers for Cloud computing" Proc. of the ACM/IEEE Conf. on Supercomputing (SC), Pittsburgh, PA, 2004, pp. 47–58.

[15]. Deepthi Dharwar, Srivatsa S. Bhat, Vaidyanathan Srinivasan, Dipankar Sarma, Pradipta Kumar Banerjee, "Approaches towards energy-efficiency in the cloud for emerging markets"in Proceedings of twenty-first ACM SIGOPS symposium on Operating systems principles. ACM, 2007, p. 278.

[16]. Francis Owusu, Colin Pattinson, "The current state of understanding of the energy efficiency of cloud computing" on 11th International Conference on Trust, Security and Privacy in Computing and Communications, IEEE, 2012, pp. 1948 –1953.

[17]. Kurt L. Keville, Rohan Garg, David J. Yates, Kapil Arya, Gene Cooperman, "Towards Fault-Tolerant Energy-Efficient High Performance Computing in the Cloud", International Conference on Cluster Computing, 2012 IEEE, pp. 622-626.

[18]. Kejiang Ye, Xiaohong Jiang, Dawei Huang, Jianhai Chen, Bei Wang "Live Migration of Multiple Virtual Machines with Resource Reservation in Cloud Computing Environments" IEEE, vol. 4, pp. 422–431, 2011.

[19]. 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 2010 IEEE/ACM International Conference on Green Computing and Communications & 2010 IEEE/ACM International Conference on Cyber, Physical and Social Computing, vol. 6161, pp. 172–182.

[20]. Liang Zhou, Baoyu Zheng, Jingwu Cui, and Sulan Tang, "Toward Green Service in Cloud: From the Perspective of Scheduling" in Proc. of the International Conference on Computing, Networking and Communications, Big Sky, MT, 2009, pp. 1–14.

[21]. Pablo Graubner, Matthias Schmidt and Bernd Freisleben, "Energy- efficient Virtual Machine Consolidation for Cloud Computing"in Information and Communication on Technology Springer, 2011, pp. 1–9.

[22]. Qi Zhang, Lu Cheng, Raouf Boutaba, "Cloud computing: state-of-the-art and research challenges" J Internet Serv Appl (2010) 1: 7–18

[23]. R.Karthikeyan , P.C hitra, "Novel Heuristics Energy Efficiency Approach for Cloud Data Center" in 2012 IEEE International Conference on Advanced Communication Control and Computing Technologies, pp. 210-216

[24]. Yamini, "Power Management in Cloud Computing Using Green Algorithm", IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012), pp.120-127.