ISSN : 2347 - 8446 (Online)
ISSN : 2347 - 9817 (Print)
Vol. 5, Issue 3 (July - Sept. 2017)
International Journal of Advanced Research in Computer Science & Technology (IJARCST 2017)

Performance Optimization with VM Provisioning Approach and Improving QoS in Cloud Computing

[I]Ramandeep Singh, [II]Jyotsna Sengupta
[I,II]Dept. of Computer Science, Punjabi University, Patiala, Punjab, India

Abstract

In Cloud Environment, the process of execution requires Resource Management due to the high process to the resource ratio. Resource Scheduling is a complicated task in cloud computing environment because there are many alternative computers with varying capacities. Scheduling is one of the most important tasks in cloud computing environment. In this paper, we have analyzed various scheduling algorithm and tabulated various parameter. We have noticed that disk space management is critical issue in virtual environment. Existing scheduling algorithm gives high throughput and cost effective but they do not consider reliability and availability. So we need algorithm that improves availability and reliability in cloud computing environment In the proposed approach the Evolutionary Algorithm is applied in this load sharing algorithm. Using the proposed approach a reserved resources are kept in order to tackle any kind of bursty traffic or bursty tasks. To validate the results of proposed approach various metrics are used like delay, delivery time, energy efficiency, throughput, down time etc. From the results section it is clear that the proposed approach is better than that of existing studies..

Keywords

Cloud, VM, Host, VM Placement Schemes

I. Introduction

With the event of high speed networks, there's associate dire rise in its usage comprised of internet queries on a daily basis and thousands of e-commerce transactions. an oversized scale knowledge centers handle this ever increasing demand by consolidating tons of and thousands of servers with different infrastructure like cooling, network systems and storage. the event of this commercialisation is called as cloud computing. Clouds square measure sky rocketing virtualized knowledge centers and applications offered as services on a subscription basis. The characteristics exhibited by Clouds square measure shown in Fig 1

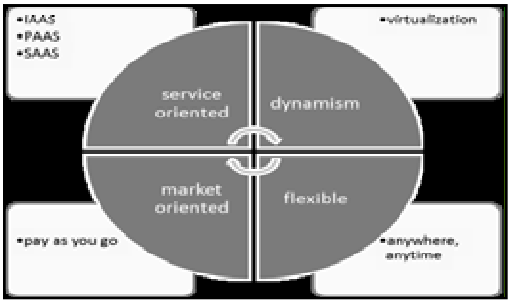

Fig 1: Characteristics of cloud computing

Cloud delivers 3 variety of services like computer code as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) [2] and might be deployed in 3 alternative ways i.e. private, public and guarded. SaaS chiefly deliver the net computer code application to the shopper, whereas PaaS offers the aptitude to form application services as on their want. It permits users to develop their computer code mistreatment programming languages and tools supported by the supplier. Infrastructure as a Service (IaaS) provides the aptitude to possess management over complete cloud infrastructure with process unit processing, storage, networks, and alternative computing resources.

Cloud may be a business model, that provides the on demand computing resources as a service to the purchasers on the rent basis. shopper have to be compelled to pay just for that quantity that's really used. To generating high revenue, cloud supplier use virtualization technology. Virtualization [5-6] is that the key technology in cloud computing, that divide the physical resources and permit the sharing of those resources. With the assistance of virtualization range of user will share an equivalent resources while not intervening to every alternative. Hypervisor may be a tiny method additionally referred to as virtual machine monitor (VMM) is employed to deploy the virtualization. It behaves like Associate in Nursing software package and accountable for taking all the choice associated with the VM. once the user demand for the pc resources (CPU, storages, network), hypervisor produce the VM and assign to the user. range of VM may be created within the single physical machine (PM). In cloud every information center keep range of host. once asking for the VM involves the hypervisor, wherever this VM is to be placed is thought as VM placement drawback. VM placement may be a NP arduous drawback [8]. thus finding an appropriate host for putting VM may be a terribly difficult task. VM placement square measure needed in 2 completely different state of affairs either for putting new VM or to position a migrated VM. Transferring the VM from one host to a different host is termed VM migration [8]. VM migration is required to manage many things like server consolidation, load equalisation, maintenances, server failure, hot spot mitigation etc. Migrations degrade the system performance, thus range of migration ought to be reduced as doable. VM is that the main process unit within the cloud. User's applications run on the VM and therefore the resource demand of those applications changes dynamically. Resources within the cloud square measure three-dimensional (CPU, memory, information measure etc.). thus resource needed by the VM could also be completely different in their quantity and kinds.

II. Smart Placement Approach

SPA is AN approach that takes under consideration the provision of the full network resource, whereas guaranteeing load-balancing and SLAs objectives. to realize this, migrating Virtual DataCenter Networks (VDNs) ought to clearly specify its elaborate resource necessities (i.e. the resource vector) to the hosting physical network. this may give for best placements and satisfying services. during this context, necessities might vary from a virtual network to a different, looking on the thought of topologies and also the provided services. However, among all the network parts, the challenge for the hosting CDNs (i.e. the physical ones) chiefly lies within the change capabilities of its network, additional exactly, its path process capacities. Indeed, wherever for a packet to urge processed through a change device, sure resources area unit needed. during this context, allow us to outline the physical switch as a collection of virtual switches, wherever every virtual switch operates a collection of virtual change ways. Mainly, a virtual change path to work needs a collection of: (1) packet process resources (network processor cycles, search caches, memories); (2) ports; (3) information measure over the ports. Typically, for a packet process task to work, this requires: (1) processors (for parsing and analysis); (2) recollections (for the search tables) which will be either internal or external (e.g. TCAMs, SRAMs); (3) queues (for packets' programing and storage, and for the method of shaping priorities); (4) information measure over the busses that interconnect the aforesaid internal parts. consequently, such physical resources are nearly divided among the various virtual knowledge ways that area unit allotted (reserved) to satisfy the wants of the VDNs topologies. Hence, for economical allocations and best placement selections, such resource vectors got to be clear so as to see for resource convenience at the hosting physical network.

III. Problem Formulation

Interesting models that tackle the problem of virtual machines placement are already proposed in the literature. However, to the best of our knowledge, no work tackles the problem of placing a full network as one package. In this context, placement models should provide for performance optimality for the whole parties involved in the placement process. Indeed, where is not a wise solution to place a virtual machine in a way that fulfills its performance requirements but causing problems to others. Considering the placement costs is also a crucial factor in such decision, however, other factors like load-balancing and QoS guarantees are also important.

• First, the candidate PMs are chosen in a way that does not consider that fact that reaching those PMs is done through a network that consists of several nodes (i.e. the connecting devices like switches and routers) interconnected via set of links (i.e. bandwidth resources). Checking the resource availability at the servers only is not enough to provide a proper candidate to hold the required applications or the needed services. Typically, this model will result in a non-optimal placement decision. Indeed, where being part of a network, it is some times more crucial to check the resource availability over the network components (e.g. switches and links) that interconnects such servers with the end-users. Without these resources, servers capacity is useless!

• Second, the migration-based placement scenario allows migrating the already placed VMs from one PM to another in a way to make space for a new VM to be place instead. This is not efficent, as those VMs that already been placed and run over a PM may have functional dependencies between each other or with other VMs from different PMs5. Migrating (re-placing) such VMs to new PMs may not be a feasible process, as this may cause several problems (e.g. service failures, interruptions) for many VMs which may violate the Service Level Agreements (SLAs) and the contracted QoS guarantees.

Third, the proposed model does not consider the load-balancing issues. Choosing the candidate PMs according to the total completion time does not necessarily comply with the load-balancing objectives. Administrators of Cloud-service networks always aim to maximize their revenue objectives while not violating the SLAs and the QoS guarantees provided to their customers. To do so, they tend to keep the loads over their networks balanced in a way to avoid congestions or any other problems that may harm the network performance.

Flow Chart

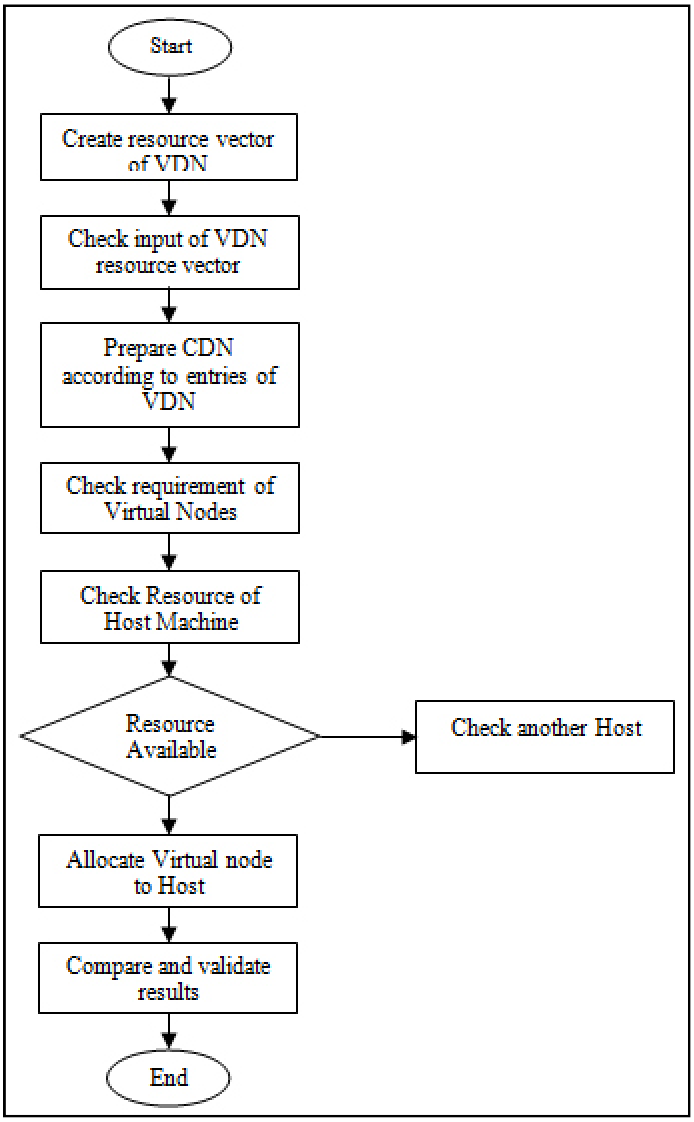

Fig 3: Flow Chart

STEP 1: Virtual machines in idle state. It describes about the main screen that is showing the machines that are in the idle state that is no load is assigned to them. In the proposed scheme the load migrated is done on the basis of parameters like utilization, speed, memory and power where VM is the virtual machine.

STEP 2: Load the machines. It describes about the machines when load is assigned to all the machines. The assigned values described about the load on various machines. When the load is allocated to the various machines continuously and it reaches the threshold value, the load will be migrated from that loaded machine to the other underloaded machine.

STEP 3: Overloaded Machine. It describes about the overloaded machine that is in the red mark. The assumed threshold for the overload condition to occur is above 80%.When the threshold is crossed, the load in the machine is migrated to the optimal

destination having less load on it.
STEP 4: According to the priority with respect to the parameters like utilization, memory, speed and power of the virtual machines, the optimal destination is chosen. It describes about the selection of the candidate Virtual Machine that to which the load is to be transferred according to the priority table. In the research a priority table is developed by the algorithm, for the calculation of the destination machine.
STEP 5: Selection of the Virtual Machine that to which the load is to be transferred according to the priority table. The optimal destination is chosen according to the priority with respect to the parameters like utilization, memory, speed and power of the virtual machines. The machine having less load on it and greater speed and better power, the load will be transferred to it. It describes about the selection of the candidate Virtual Machine that to which the load is to be transferred according to the priority table.
STEP 6: Downtime during the load sharing. Downtime is defined as the time at which the virtual machines stop executing. It includes transfer of the processor state. In the proposed approach, the downtime is decreased which results in better performance. The downtime can be calculated by the formula:
Total Downtime = Stop-and-copy + commitment + activation.
STEP 7: Efficiency. In the proposed approach, the efficiency of the system is improved.

V. Results and Discussion

Make Span: In the proposed approach, the makespan is decreased which results in better performance.

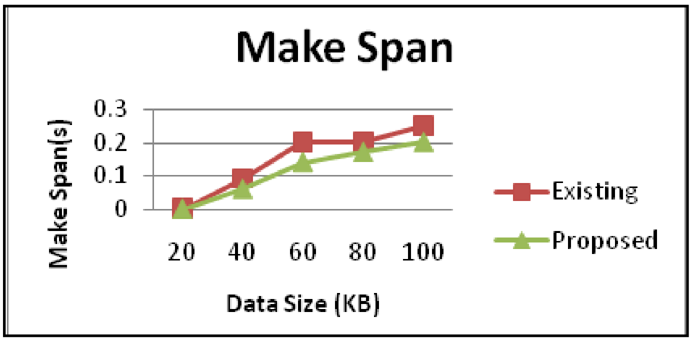

Fig 4: Makespan

Load: In the proposed approach, the load of the system is improved.

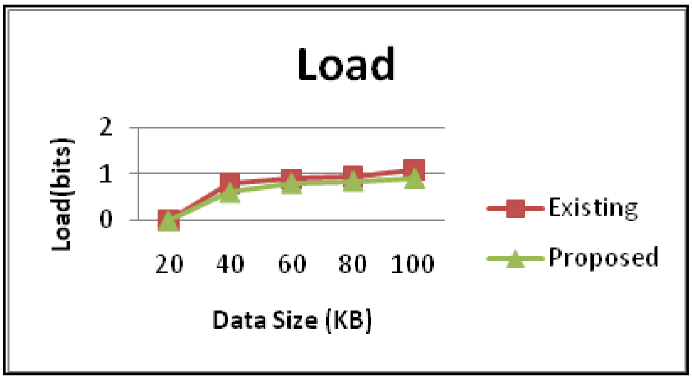

Fig 5: Load

Throughput: Throughput may be defined as the ratio of output to input in a system. In this graph the throughput of the proposed system is better than that of existing scheme.

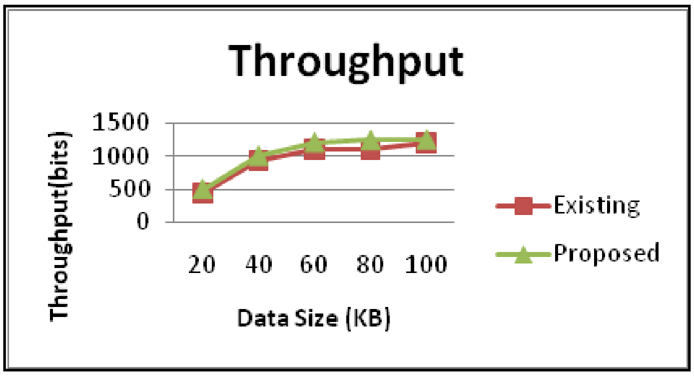

Fig 6: Throughput

Average Energy Consumption: It may be defined as the power that is consumed to perform certain operations in a virtual machine. From the graph it is very clear that the proposed system is more efficient than that of existing one.

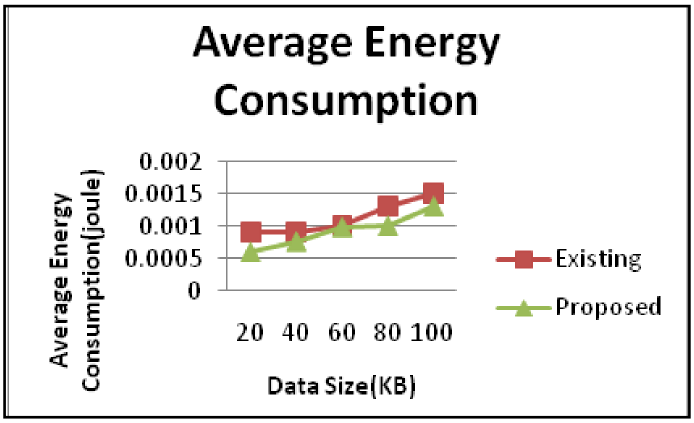

Fig 7: Average Energy Consumption

Comparison Table of work performed:
Here the comparison takes place between the base paper and the work performed. The results produced by the work are better than the previous work done.

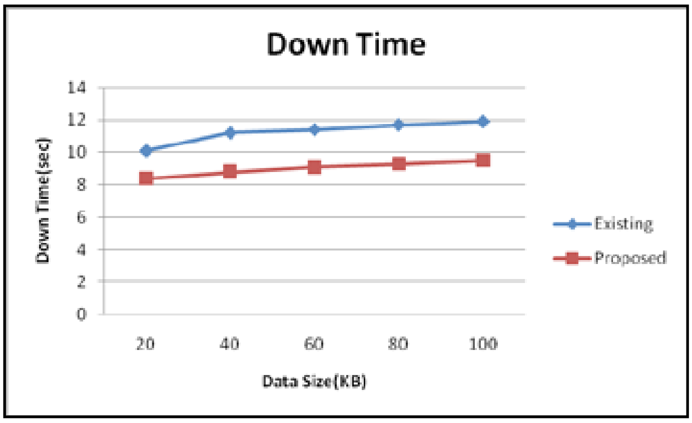

Fig 8: Down Time(sec)

Down Time: This may be defined as the time span for which the system was in the idle state, in which no work is done. From the graph it is clear that the down time in proposed system is less than that of existing system.

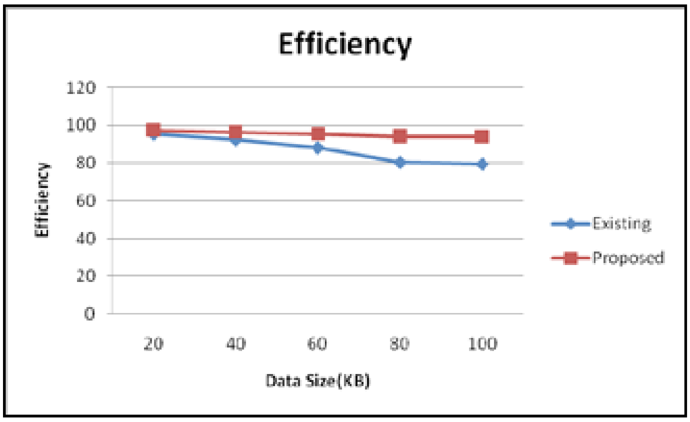

Fig 9: Efficiency

Efficiency: This may be defined as the ratio of output to the input. Efficiency of a system may define the performance of it. From the graph it is clear that the efficiency in proposed system is more than that of existing system.

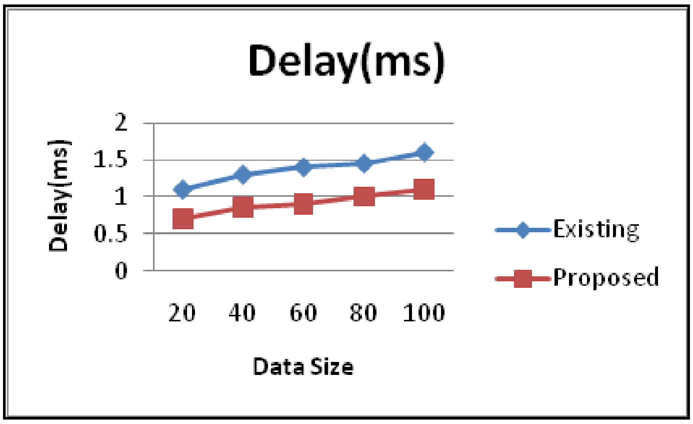

Fig 10: Delay

Delay: This may be defined as the time taken for a task to be completed. From the graph it is clear that the delay in proposed system is less than that of existing system.

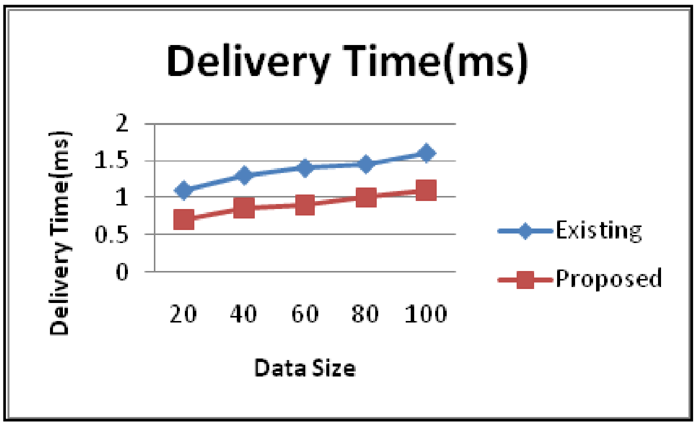

Fig 11: Delivery Time (ms)

Delivery Time(ms): This may be defined as the time taken for a task to be delivered. From the graph it is clear that the delay in proposed system is less than that of existing system.

Table 1: Comparison table of work performed

Approach	Down Time(sec)	Efficiency(%)	Make Span(ms)	Load(bits)	Throughput (bits)	Avg Energy Consumption (joule)	Delay	Delivery Time
SPA	11.3	85	0.2	0.9	1100	0.0013	10.6	9.5
Improved SPA	9.7	93	0.14	0.8	1200	0.001	8.7	7.9

Downtime is defined as the time at which the virtual machines stop executing. It includes transfer of the processor state. In the proposed approach, the downtime is decreased which results in better performance.

VI. Conclusion

This implementation aims towards the establishment of performance qualitative analysis on load sharing in VM to VM and then implemented in CloudSim with Java language. Here major stress is given on the study of load balancing algorithm with heterogeneous resources of the cloud, followed by comparative survey of other algorithms in cloud computing with respect to scalability, homogeneity or heterogeneity and process migration. A previous study also indicates change of MIPS will affect the response time and increase in MIPS versus VM decreases the response time. When image size of VM is implemented against the VM bandwidth then no significant effect is found on response time and it remains constant for which these parameters are investigated. But in case of Cloudlet long length versus Host bandwidth a pattern is observed in which response time increases in proportionate manner. Using the modified approach the reduction in the down time of the various processes are achieved as shown in results.

References

[1]. *Ahmad Nahar Quttoum, Mohannad Tomar, Bayan Khawaldeh, Rana Refai, Alaa Halawani, Ahmad Freej, "SPA: Smart Placement Approach for Cloud-service Datacenter Networks", The 10th International Conference on Future Networks and Communications, Vol: 56, 2015, pp: 341-348*

[2]. *Chonglin Gu, Pengzhou Shi, Shuai Shi, Hejiao Huang, Xiaohua Jia, "A Tree Regression-Based Approach for VM Power Metering", Special Section On Big Data For Green Communications And Computing, IEEE, June 1, 2015*

[3]. *Zhaoning Zhang, Ziyang Li, Kui Wu, Dongsheng Li, Huiba Li, Yuxing Peng, Xicheng Lu, "VMThunder: Fast Provisioning of Large-Scale Virtual Machine*

Clusters", IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, Vol. 25, No. 12, DECEMBER 2014

[4]. *Jiaxin Li, Dongsheng Li, Yuming Ye, Xicheng Lu, "Efficient Multi-Tenant Virtual Machine Allocation in Cloud Data Centers", TSINGHUA SCIENCE AND TECHNOLOGY, ISSN: 1007-0214, Volume 20, Number 1, February 2015, pp: 81-89*

[5]. *Gursharan Singh, Sunny Behal, Monal Taneja, "Advanced Memory Reusing Mechanism for Virtual Machines in Cloud Computing", 3rd International Conference on Recent Trends in Computing, Vol: 57, 2015, pp: 91-103*

[6]. *Narander Kumar, Swati Saxena, "Migration Performance of Cloud Applications- A Quantitative Analysis", International Conference on Advanced Computing Technologies and Applications, Vol. 45, 2015, pp: 823-831*

[7]. *Aarti Singh, Dimple Juneja, Manisha Malhotra, "Autonomous Agent Based Load Balancing Algorithm in Cloud Computing", International Conference on Advanced Computing Technologies and Applications, Vol: 45, 2015, pp: 832-841*

[8]. *S. Sohrabi, I. Moser, "The Effects of Hotspot Detection and Virtual Machine Migration Policies on Energy Consumption and Service Levels in the Cloud", ICCS, Vol: 51, 2015, pp: 2794-2798*

[9]. *Mohammad Mehedi Hassan, Atif Alamri, "Virtual Machine Resource Allocation for Multimedia Cloud: A Nash Bargaining Approach", International Symposium on Emerging Inter-networks, Communication and Mobility, Vol: 34, 2015, pp: 571-576*

[10]. *Christina Terese Josepha, Chandrasekaran K, Robin Cyriaca, "A Novel Family Genetic Approach for Virtual Machine Allocation", International Conference on Information and Communication Technologies, Vol: 46, 2015*

[11]. *Antony Thomas, Krishnalal G, Jagathy Raj V P, "Credit Based Scheduling Algorithm in Cloud Computing Environment", International Conference on Information and Communication Technologies, Vol: 46, 2015, pp: 913-920*

[12]. *Doshi Chintan Ketankumar, Gaurav Verma, K. Chandrasekaran, "A Green Mechanism Design Approach to Automate Resource Procurement in Cloud", Eleventh International Multi-Conference on Information Processing, Vol: 54, 2015, pp: 108-117*

[13]. *A.I.Awad, N.A.El-Hefnawy, H.M.Abdel_kader, "Enhanced Particle Swarm Optimization For Task Scheduling In Cloud Computing Environments", International Conference on Communication, Management and Information Technology, Vol: 65, 2015, pp: 920-929*

[14]. *Arunkumar. G, Neelanarayanan Venkataraman, "A Novel Approach to Address Interoperability Concern in Cloud Computing", 2nd International Symposium on Big Data and Cloud Computing, Vol: 50, 2015, pp: 554-559*

[15]. *Atul Vikas Lakra, Dharmendra Kumar Yadav, "Multi-Objective Tasks Scheduling Algorithm for Cloud Computing Throughput Optimization", International Conference on Intelligent Computing, Communication & Convergence, Vol: 48, 2015, pp: 107-113*

[16]. *Narander Kumar, Swati Saxena, "A Preference-based Resource Allocation In Cloud Computing Systems", 3rd International Conference on Recent Trends in Computing, Vol: 57, 2015, pp: 104-111*

[17]. *Nidhi Bansal, Amitab Maurya, Tarun Kumar, Manzeet Singh, Shruti Bansal, "Cost performance of QoS Driven task scheduling in cloud computing", Third International Conference on Recent Trends in Computing, Vol: 57, 2015, pp: 126-130*

[18]. *Mehmet Sahinoglua, Sharmila Ashokan, Preethi Vasudev, "Cost-Efficient Risk Management with Reserve Repair Crew Planning in CLOUD Computing", Cost-Efficient Risk Management with Reserve Repair Crew Planning in CLOUD Computing, Vol: 62, 2015, pp: 335-342*

[19]. *Patricia Arroba, Jos´e L. Risco-Mart´ın, Marina Zapater, Jos´e M. Moya, Jos´e L. Ayala, Katzalin Olcoz, "Server Power Modeling for Run-time Energy Optimization of Cloud Computing Facilities", 6th International Conference on Sustainability in Energy and Buildings, Vol: 62, pp: 2014, pp: 401-410*