IPHV7I20063X
International Journal Of Advanced Research and Innovation -Vol.7, Issue .II
ISSN Online: 2319 – 9253
Print: 2319 – 9245

PRIORITY BASED RESOURCE ALLOCATION STRATAGIES IN CLOUD COMPUTING

[#1]KOTAPATI UPENDRA - M.Tech Pursuing,
[#2] V. Chandrakala - Assistant Professor,
Department of Computer Science and Engineering,
KKR AND KSR INSTITUTE OF TECHNOLOGY AND SCIENCES (KITS) GUNTUR.

Abstract: *Cloud computing has emerged as the default paradigm for a variety of fields especially considering the resources and infrastructures consumption in case of distributed access. The solution has however placed a lot of emphasis of a cloud server with variety of demands of which quality of service reminds a paramount strategy. There are a lot of strategies is in place for this quality cost which is regulated by service level agreement [SLA].An SLA is an agreement between client and server which when violated will impose penalties for the infringement or violation performance evaluation place a key role allowing system managers to evaluate the effects of different resource management strategies on the data center functioning and to predict the corresponding cost/benefits. In order to deal with very large systems composed of hundreds or thousands of resources. The system should allow to easily implementing different strategies and should have policies to represent different working conditions. Keeping this is mind a rewards scheme called as SRNS-Stochastic Reward Nets is utilized which is dynamic in nature to the status of the requests made and job allocated. The proposed model is scalable enough to represent system composed of thousands of resources and it makes possible to represent both physical and virtual resources exploiting cloud specific concepts such as the infrastructure elasticity.*

Keywords: *Cloud Computing, Resource Allocation, Static, Dynamic, Cloud-Oriented Performance Metrics, Resiliency, Responsiveness, Stochastic Reward Nets.*

I.INTRODUCTION

Cloud Computing is the emerging technology where the services are provided over the Internet. This has become a preferred option for many business contexts. Rather than owning the resources, they are used on a rented basis. This avoids the cloud user the head ache of managing the resources which has to be owned by them instead. For providing the service by the cloud service provider and using the services by the cloud user, an initial agreement called the Service Level Agreement (SLA) has to be made between the cloud users and the cloud service provider. While resource allocation is made to the cloud user, SLA violation should be avoided as much as possible or SLA violation should be minimal without compromising Quality of Service (QoS) parameters like performance, availability, response time, security, reliability, throughput etc. In order to efficiently utilize resources, various resource allocation techniques are to be used. For allocating the resources, virtualization technique has to be used. Since the actual physical resources available are less than the resource demand, virtual resources are to be created, mapped onto the physical resources, and these virtual resources are to be allocated to the requesting cloud user. After utilization of the virtual resources, they have to be destroyed by the cloud service provider. All these have to be made without violating SLA by meeting QoS.

With the limited amount of physical resources available, resource allocation becomes a challenging task for the cloud service provider. Since cloud computing is a multi-tenancy model, multiple users‟ requests for the cloud resources. So cloud service provider has to decide on how many virtual resources are to be created based on the cloud users‟ requests. Also which virtual machine (VM) has to be mapped onto which physical machine (PM) has to be taken care. That is, VM-PM mapping techniques have to be considered. At what instance VM migration has to be done is also based on identifying heavily loaded node and lightly loaded node. So the ultimate goal of the cloud service provider is to maximize profit and maximize resource utilization and the goal of cloud user is to minimize payment by renting the resources.

The cloud computing model is comprised of a front end and a back end. These two elements are connected through a network. The front end is the vehicle by which the user interacts with the system and the back end is the cloud itself. The front end is composed of a client computer, or the computer network of an enterprise, and the applications used to access the cloud. The back end provides the applications, computers, servers, and data storage that creates the cloud of services. Cloud computing describes a type of outsourcing of computer services, similar to the way in which the supply of electricity is outsourced. Users can simply use it. They do not need to worry where the electricity is from, how it is produced, or transported. In cloud, services allowing users to easily access resources anywhere anytime. Users can pay for a service and access the resources made available during their subscriptions until the subscribed periods expire. Users are then forced to demand such resources if they want to access them also after the subscribed periods. IaaS providers build flexible cloud solutions according to the hardware requirements of customers; furthermore it let customers run operating systems and software applications on virtual machine (VMs).Customers merely pay for the resources that are actually used.

There are various parameters to be considered while allocating resources. While allocating resources to the cloud user, underutilization (wastage) of resources due to over provisioning and overutilization (due to under provisioning) should be avoided. Allocation of resources should consider various parameters like Quality of Service parameters like response time,performance, availability, reliability, security, throughput etc.

Performance: For some application demands, performance is one of the important criteria. The system should perform well to provide service to the cloud user.

Response Time: For interactive applications, response time is an important factor. The system must respond well for these kinds of applications. Reliability: The system used should be reliable so that the cloud user has no head ache of changing the system.

Availability: Whenever cloud resources are requested the cloud service provider must be able to allocate within short span.

Security: For critical applications like online transaction applications, system used has to be secure. Otherwise it is not safe to use such a kind of system.

Throughput: No. of applications run per unit time should be more.

For efficient resource allocation all the above parameters are to be considered.

The rest of the paper is organized as follows. Section II discusses Static Allocation Techniques, Section III discusses Dynamic Allocation Techniques, Section IV gives a comparative study of resource allocation techniques and Section Vgives conclusion.

There are two major types of resource allocation techniques.

Static Allocation: In this case, the cloud user has to make prior request for the resources. By doing so user knows what resources are required and how many instances of the resources are needed ahead of using the system. But the drawback in this case is it leads to underutilization or overutilization of resources depending on the time the application is run.

Dynamic Allocation: Cloud resources are requested by the cloud user on the fly or as and when the application needs. Here underutilization and overutilization of resources is avoided as much as possible. But the requested resources might not be available when requested on the fly. The service provider has to make allocationfrom other participating cloud data center.

Resource Allocation Techniques: There are many resource allocation techniques both static and dynamic each having its own advantages and limitations.

II. STATIC RESOURCE ALLOCATION TECHNIQUES

Ruben Van den Bossche et. al. [10] have discussed about online cost-efficient scheduling. For efficient utilization of resources and to minimize cost, a hybrid cloud scheduling algorithm is used for hybrid clouds. The need for hybrid cloud determines which workload type has to be outsourced and to which cloud service provider. These decisions should minimize the cost of running the part of the total workload on one or more multiple cloud providers which must take into account application requirement such as deadline constraint and data requirement. Variety of cost factors, pricing models, and cloud provider offerings further makes use of automated scheduling approach in hybrid clouds. These issues are tackled by a set of algorithms to cost efficiently schedule the deadline constrained set of applications on both public cloud provider and private cloud infrastructure. These algorithms take into account both computational and data transfer costs as well as network bandwidth constraints.

Siva Theja Maguluri et. al. [26] have discussed about Heavy Traffic Optimal Resource Allocation Algorithms. In this paper, a stochastic model of cloud computing is studied, where jobs arrive according to a stochastic process and request resources like CPU, memory and storage space. A model is considered where the resource allocation problem can be separated into a routing or load balancing problem and a scheduling problem. Join-the-shortest-queue routing and power of-two-choices routing algorithms with MaxWeight scheduling algorithm is studied. These algorithms are throughput optimal and it shows that these algorithms are queue length optimal in the heavy traffic limit.

Mohammad Firoj Mithani et. al. [17] have discussed about Resource Allocation in Multi-Tier Cloud Systems. The problem of over provisioning of resources for enterprise-class applications hosted in cloud systems is addressed and ways to minimize over-provisioning of IT resources in multi-tier cloud systems is proposed by adopting an innovative approach of application performance monitoring and resource allocation at individual tier levels, on the basis of criticality of the business services and availability of the resources. To prevent over-provisioning of resources and improve the cloud managed resource allocation processes, an intermediate „Tier-centric Business Impact and Cost Analysis‟(T-BICA) capability to monitor performance of the business application tiers, including infrastructure components at each tier by leveraging existing monitoring layer is introduced. Ikki Fujiwara et. al. [25] have applied double-sided combinatorial auctions to resource allocation. A market-based resource allocation will be effective in a cloud computing environment where resources are virtualized and delivered to users as services. Such a market mechanism is proposed to allocate services to participants efficiently. The mechanism enables users (1) to order a combination of services for workflows and co allocations and (2) to reserve future/current services in a forward/spot market. The forward market and the spot market run independently to make predictable and flexible allocations at the same time.

Davide Adami et. al. [4] have proposed effective control strategies using openflow in Cloud Data Center. A unified control and management of computing and network resources would be required for assuring proper traffic performances in high volatile virtual machine deployments. This paper introduces a novel resource control platform for virtualized DC environments aimed at optimizing virtual machine placement on physical servers also considering traffic load across links in order to limit oversubscription-related problems. For this purpose Open Flow statistics are elaborated for their distinguished features that well suit virtualized environments. Two novel algorithms have been proposed and the effectiveness of proposed traffic-aware VM placement strategies is evaluated.

Hong Xu et. al. [2] have proposed Anchor, a versatile and efficient framework for Resource management. Anchor is a general resource management architecture that uses the stable matching framework to separate policies from mechanisms when mapping virtual machines to physical servers. In Anchor, clients and operators are able to express a variety of different resource management policies, and these policies are captured as preferences in the stable matching framework. The highlight of Anchor is a new many-to-one stable matching theory that efficiently matches VMs with heterogeneous resource needs to servers, using both offline and online algorithms.

Sheng Di et. al. [5] have discussed about error-tolerant allocation and payment minimization. With virtual machine (VM) technology being increasingly mature, compute resources in cloud systems can be partitioned in fine granularity and allocated on demand. Three contributions in this paper are: 1) Deadline-driven resource allocationproblem is formulated based on the cloud environment facilitated with VM resource isolation technology, and also a novel solution with polynomial time is proposed, which could minimize users‟ payment in terms of their expected deadlines. 2) By analyzing the upper bound of task execution length based on the possibly inaccurate workload prediction, further an error-tolerant method is proposed to guarantee task‟s completion within its deadline. 3) Its effectiveness over a real VM-facilitated cluster environment under different levels of competition is validated.

Shuo Zhang et. al. [9] have proposed a novelresource allocation algorithm for a Heterogeneous Data Center. The issue concerned about data centers is huge energy consumption, which not only increases operational cost, but also causes environment problems. One way to reduce energy consumption is decreasing the scale of data centers, and another way is using well-designed resource allocation algorithm to achieve the tradeoff between energy consumption and performance. This paper focuses on well-designed resource allocation algorithm. First the resource allocation problem in a heterogeneous data center is discussed, and then a novel resource allocation algorithm is proposed based on fair scheduling policy and energy-aware policy to this problem.

Ping Guo et. al. [6] have discussed about hierarchical resource management model. In order to resolve the heavy load and the long waiting time in the centralized resource management, new better model, which is the hierarchical resource management model based on resource tables is proposed. This paper describes the model of hierarchical management system and the process of the resource allocation; giving the form of the resource table in this model and the access controlpolicy of the resource table.

III. DYNAMIC RESOURCE ALLOCATION TECHNIQUES

Valérie Danielle Justafort et. al. [13] discusses about Performance-Aware Virtual Machine Allocation. Cloud Computing has major challenges among which virtual machine (VM) placement is considered as one of the crucial problems. From the cloud provider‟s perspective, this process often consists in choosing the best convenient physical hosts with regards to energy efficiency, resource utilization and revenues maximization. However, to secure the loyalty of their clients, cloud providers should also consider the applications performance. In this paper, VM placement approach is proposed tackling the applications performance concern, in an inter cloud environment. VM placement is stated as a mixed integer programming problem which aims to maximize a global utility function considering VM bandwidth requirements and network latencies.

Bo Yin et. al. [22] discusses about multi-dimensional resource allocation. In this paper, study on the resource allocation at the application level is done, instead to map the physical resources to virtual resources for better resource utilization in cloud computing environment. A multi-dimensional resource allocation (MDRA) scheme for cloud computing is proposed that dynamically allocates the virtual resources among the cloud computing applications to reduce cost by using fewer nodes to process applications. In this model, a two-stage algorithm is adopted to solve this multi-constraint integer programming problem. The algorithm can dynamically reconfigure the virtual machines for cloud applications according to the load changes in cloud applications by assigning new applications on the working node instead of opening a new node. Experiment results show that this algorithm can save resources and increase resource utilization as well as centralize working nodes. MDRA provide resource to deal with requirements more steadily. Moreover, the proposed algorithm in the long run can save power efficiently when the demand of user is decreasing. Amit Nathani et. al. [14] have proposed policy based resource allocation in IaaS cloud. Most of the Infrastructure as a Service Cloud (IaaS) uses simple resource allocation techniques like immediate and best effort. Immediate Resource Allocation Policy allocates the resource if available otherwise the request is rejected. Best effort policy also resources if currently available otherwise the request is placed in the FIFO queue. So it is not possible for the Service provider to satisfy all the requests as there is finite number of resources at a time. So Haizea is a resource lease manager that addresses these issues by using complex resource allocation strategies. Haizea uses resource leases as resource abstraction and implements these leases by allocatingVirtual Machines (VMs). Haizea supports four resource allocation policies: Immediate, best effort, advanced reservation and deadline sensitive. Among the four, deadline sensitive leases by Haizea supports minimal rejection of leases and using this policy, maximum resource utilization is possible. Dynamic Planning based scheduling algorithm is implemented in Haizea which can admit leases and prepare a schedule whenever a new lease can be accommodated. Sharrukh Zaman et. al. [1] discusses online mechanism for dynamic VM provisioning and allocation. Current cloud provider allocates virtual machine instances via fixed price-based or auction based mechanisms. The limitation in these mechanisms is they are all offline mechanism and they need to collect information and invoked periodically.This limitation is addressedby designing an online mechanism by dynamically provisioning and allocating Virtual machine Instances in Clouds. The algorithm called Mechanism for online VM provisioning and allocation (MOVMPA) is invoked as soon as the user makes a request or some VM instances already become available again. When invoked, the mechanism selects users who would be allocated VM instances they requested for and ensures that those users will continue using for the entire period they requested for.

IV.PROPOSED MODEL

The proposed Model presents a stochastic model, based on Stochastic Reward Nets (SRNs), that exhibits the above mentioned features allowing capturing the key concepts of an IaaS cloud system. The proposed model is scalable enough to represent systems composed of thousands of resources and it makes possible to represent both physical and virtual resources exploiting cloud specific concepts such as the infrastructure elasticity. Cloud based systems are inherently large scale, distributed, almost always virtualized, and operate in automated shared environments. Performance and availability of such systems are affected by a large number o parameters including characteristics of the physical infrastructure (e.g., number of servers, number of cores per server, amount of RAM and local storage per server, configuration of physical servers, network configuration, persistent storage configuration), characteristics of the virtualization infrastructure (e.g., VM placement and VM resource allocation, deployment and runtime overheads),failure characteristics (e.g., failure rates, repair rates, modes of recovery),characteristics of automation tools used tom manage the cloud system, and so on. Because of this, any naïve modeling approach will quickly run into state explosion and/or intractable solution. The proposed model is scalable enough to represent systems composed of thousands of resources and it makes possible to represent both physical and virtual resources exploiting cloud specific concepts such as the infrastructure elasticity. With respect to the existing literature, the innovative aspect of the present work is that a generic and comprehensive view of a cloud system is presented. Low level details, such as VM Multiplexing, are easily integrated with cloud based actions such as federation, allowing investigating different mixed strategies. An exhaustive set of performance metrics are defined regarding both the system provider (e.g., utilization) and the final users (e.g., responsiveness).

4.1 CLOUD MODEL AND RESOURCES

Clouds are modeled in the server client mode by a datacenter component for handling service requests. These requests are application elements sandboxed within VMs, which need to be allocated a share of processing power on Datacenter’s host components. By VM processing, it means that a set of operations related to VM life cycle: provisioning of a host to a VM, VM creation, VM destruction, and VM migration. A Datacenter is composed by set hosts, which are responsible for managing VMs during their life cycles. Host is a component that represents a physical computing node in a cloud: it is assigned a pre-configured processing capability (expressed in million of instruction per second-MIPS), memory, storage, and a scheduling policy for allocating processing cores to virtual machines. The Host component implements interfaces that support modeling and simulation of both single-core and multi-core nodes.

4.2 CLIENT REQUESTS

Clients are registered with the server for utilizing the resources and application. Needs to access the resources and applications are implemented and customary SLA’s are put in place to handle the request handling mechanisms. Each Request is considered as a job to be completed. A job is usage of a app or a resource or both and sending back the processed requested data. The job arrival or rather the request process constituter three different scenarios. In the first one (Constant arrival process) the arrival process be a homogeneous Poisson process with uniform rate. However, in large scale distributed systems with thousands of users, such as cloud systems, could exhibit self-similarity/long range dependence with respect to the arrival process and for these reasons, in order to take into account the dependencies of the job arrival rate on both the days of a week, and the hours of a day, in the second scenario is the periodic arrival process. This is also chosen to model the job arrival process. This is also chosen to model the job arrival process as Markov Modulated Poisson Process (MMPP).

4.3 CLOUD FEDERATION AND MONITORING

This is the analytical algorithm part, where parameters like weight; intermediate requests like waiting time, bandwidth calculation completion time are all executed here. This is based on the inputs the optimal solution is arrived federation with other clouds is modeled allowing tokens in place p_{queue} to be moved, through transition t_{upload}, in the upload queue represented by place p_{send}. In accordance with the assumptions made before, transition t_{upload} is enabled only if the number of tokens in place p_{queue} is greater than Q and the number of tokens in place P_{send} is less than D. Moreover, in order to take into account the federated cloud availability, concurrent enabled transition t_{upload} and t_{drop} are managed by setting their weights.

4.4 VIRTUAL MULTIPLEXING

VM, whose management during its life cycle is the responsibility of the host component. As discussed earlier, a host can simultaneously instantiate multiple VMs and allocate cores

IPHV7I20063X
International Journal Of Advanced Research and Innovation -Vol.7, Issue .II
ISSN Online: 2319 – 9253
Print: 2319 – 9245

based on predefined processor sharing policies(space-shared, time-shared).Every VM component has access to a component that stores the characteristics related to a VM's internal scheduling policy, which is extended from the abstract component called VM Scheduling. The model supports VM scheduling at two levels: First, at the host level and second, at the VM level. At the host level, it is possible to specify how much of the overall processing power of each core in a host will be assigned to each VM. At the VM level, the VM assign specific amount of the available processing power to the individual task units that are hosted within its execution engine.

FIGURES

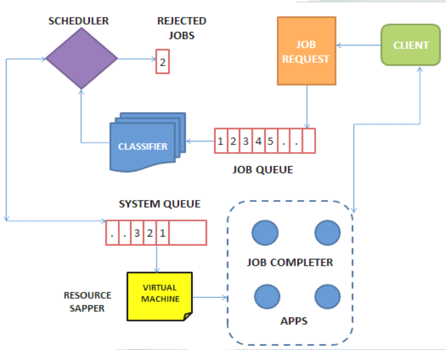

Fig 3.1 Scheduling Process

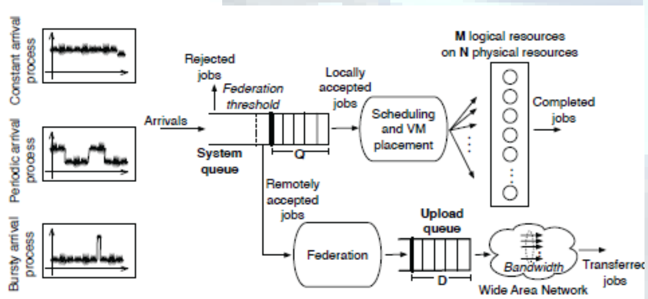

Fig3.2 System Architecture

V. COMPARISON OF RESOURCE ALLOCATION APPROACHES

Table 1 gives comparative study of Static Resource Allocation techniques. Table 2 gives a comparative study of Dynamic Resource Allocation Techniques. Table 3 gives a comparative study of Resource Allocation Techniques based on other parameters:

Table 1. Comparison of Static resource Allocation techniques

Allocation Strategies	Revenue Maximization	Efficient Resource Utilization	Improve Performance	Fair Scheduling	Efficient VM-PM Mapping	Cost Minimization
Hybrid Cloud Scheduling	✓	✓				✓
Tier-Centric Business Impact and Cost Analysis approach		✓		✓		
Anchor Resource Management Framework		✓	✓	✓	✓	
Efficient resource Control Strategies				✓	✓	
Error Tolerant resource Allocation						✓
Hierarchical Resource Management Model		✓				
Novel Resource Allocation				✓		
Multilayer Resource Management		✓	✓			
Berger Model based Job Scheduling				✓		
Heavy Traffic Optimal Resource Allocation			✓	✓		
Double-sided Combinational Auctions	✓			✓		
Security-Aware Resource Allocation				✓		

IPHV7I20063X
International Journal Of Advanced Research and Innovation -Vol.7, Issue .II
ISSN Online: 2319 – 9253
Print: 2319 – 9245

Table 2.Comparison of Dynamic Resource Allocation Techniques

Allocation Strategies	Revenue Maximization	Efficient Resource Utilization	Improve Performance	Fair Scheduling	Efficient VM-PM Mapping	Cost Minimization
Policy based Resource Allocation		✓				
Online Mechanism for Dynamic VM Provisioning and Allocation		✓				
Resource Reconstruction Algorithm		✓				
Priority-based Scheduling	✓	✓				
Performance Aware VM Allocation			✓			
Minimum cost maximum Flow Resource Allocation	✓	✓	✓			✓
VM migration based Resource Management					✓	
Overbooking based Resource Allocation	✓	✓	✓			
Optimized resource Allocation	✓	✓	✓			✓
Location-Aware Dynamic Resource Allocation		✓	✓			
Agent-based Adaptive Resource Allocation		✓	✓			
Multi-dimensional Resource Allocation		✓				

Table 3: Comparison of Resource Allocation techniques based on other parameters

Allocation Strategies	Fault tolerant	Prevent Over provisioning	Security	Energy Aware Policy
Error Tolerant resource Allocation	✓			
Novel Resource Allocation				✓
Security Aware Resource Allocation			✓	
Overbooking based resource allocation		✓	✓	
Multidimensional resource Allocation				✓
Tier-Centric Business Impact and Cost Analysis approach		✓		

VI. CONCLUSION:

In this paper, SRNs allow us to define reward functions that can be associated to a particular state of the model in order to evaluate the performance level reached by the system during the sojourn in order in that state performance metrics able to characterize the system behavior from both the provider and the user point-of-views.the stochastic model to evaluate the performance of an IaaS cloud system. Several performance metrics have been defined, such as availability, utilization, and responsiveness, allowing to investigate the impact of different strategies on both provider and user point-of-views. So it becomes challenging for the cloud service provider to allocate resources considering various parameters like performance, availability, revenue maximization, reliability, security, response time, payment minimization, energy efficiency etc. depending on the application needs. The above algorithms satisfy some of the parameters even though all cannot be satisfied. . In future the model can be implemented to include intra and reservoir clouds where other performance parameters like application takes sufficient memory and bandwidth issues are considered. This model may be replicated in web services domain where similar features are used and service agreements are made between the website client and web service provider so the concept can be extended to cover such models and issues as well in the future.

REFERENCES:

1] Sharrukh Zaman, Daniel Grosu, "An Online Mechanism for Dynamic VM Provisioning and Allocation in Clouds", IEEE Fifth International Conference on Cloud Computing, 24-29 June 2012, pp 253-260, DOI 10.1109/CLOUD.2012.26.

2] Hong Xu, Baochun Li, "Anchor: A Versatile and Efficient Framework for Resource Management in the Cloud", IEEE Transactions On Parallel And Distributed Systems, Vol. 24, No. 6, June 2013, pp 1066-1076, DOI 10.1109/TPDS.2012.308.

3] Mayank Mishra, Anwesha Das, Purushottam Kulkarni, Anirudha Sahoo, "Dynamic Resource Management Using Virtual Machine Migrations", IEEE Communications Magazine, Vol 50, Issue 9, September 2012, DOI 10.1109/MCOM.2012.6295709.

4] Xiao-Jun Chen, Jing Zhang, Jun-Huai Li, Xiang Li, "Resource Reconstruction Algorithms for On-demand Allocation in Virtual Computing Resource Pool", Springer, International Journal of Automation and Computing, Vol 9, Issue 2, pp 142-154, DOI: 10.1007/s11633-012-0627-3.

5] Mohammad Firoj Mithani, Shrisha Rao, "Improving Resource Allocation in Multi-Tier Cloud Systems", IEEE International Conference on Systems Conference, 19-22 March 2012, pp 1-6, DOI: 10.1109/SysCon.2012.6189471.

6] V.Vinothina, R.Sridaran, Padmavathi Ganapathi, "A Survey on Resource Allocation Strategies in Cloud Computing", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 3, No.6, 2012, pp 97-104.

7] Ronak Patel, Sanjay Patel, "A Survey on Resource Allocation Strategies in Cloud Computing", International Journal of Engineering Research & Technology (IJERT), Vol.2, Issue 2, Feb 2013.

8] N.Asha, G. Raghavendra Rao, "A Review on Various Resource Allocation Strategies in Cloud Computing", International Journal of Emerging Technology and Advanced Engineering (IJETAE), Vol.3, Issue 7, July 2013.

9] Bo Yin, Ying Wang, Luoming Meng, Xuesong Qiu "A Multi-dimensional Resource Allocation Algorithm in Cloud Computing", Journal of Information & Computational Science 9: 11 (2012) 3021–3028.

10] Gihun Jung, Kwang Mong Sim ,"Agent-based Adaptive Resource Allocation on the Cloud Computing Environment", 2011 International Conference on Parallel Processing Workshops, pp 345-351, DOI: 10.1109/ICPPW.2011.18.

11] Gihun Jung, Kwang Mong Sim, "Location-Aware Dynamic Resource Allocation Model for Cloud Computing Environment", International Conference on Information and Computer Applications, IPCSIT vol. 24, 2012, pp 37-41.

12] Ikki Fujiwara, Kento Aida, Isao Ono, "Applying Double-sided Combinational Auctions to Resource Allocation in Cloud Computing", IEEE, pp 7-14, 2010, DOI 10.1109/SAINT. 2010.93.

13] Siva Theja Maguluri, R. Srikant, Lei Ying,"Heavy Traffic Optimal Resource Allocation Algorithms for Cloud Computing Clusters", 24th International Teletraffic Congress, 2012, Article No. 25. 27] Saeed Al-Haj, Ehab Al-Shaer, HariGovind V.

14] Ramasamy "Security-Aware Resource Allocation in Clouds", IEEE 10th International Conference on Services Computing, Santa Clara, CA, June 28 2013-July 3 2013, pp 400 – 407, DOI: 10.1109/SCC.2013.36.

[15] R. Buyya et al., "Cloud computing an emerging it platforms: Vision, hype, and reality for delivering computing as the 5th utility , "Future Gener, comput. Syst.vol.25 pp.599-616,june 2009.

[16] H. Liu et al., "Live virtual machine migration via asynchronous replication and state synchronization ," parallel and distributed Systems, IEEE Transactions on, vol.22,no.12 , pp.1986-1999,dec.2011.

[17] B. Rochwerger et al., "Reservoir – when one cloud is not enough," Computer, vol.44 , no.3, pp.44-55,march 2011.