Secure Dynamic Resource Provisioning Cost by Optimized Placement of Virtual Machines in Cloud Computing

Aarthee.S, Venkatesan.R
PG Scholar, Department of CSE, Sudharsan Engineering College, Pudukkottai, TamilNadu, India
aartheerenu@gmail.com
Assistant Professor, Department of CSE, Sudharsan Engineering College, Pudukkottai, TamilNadu, India
venkatesaneng@gmail.com

Abstract - Cloud computing provides pay-as-you-go computing resources and accessing services are offered from data centers all over the world as the cloud. Consumers may find that cloud computing allows them to reduce the cost of information management as they are not required to own their servers and can use capacity leased from third parties or cloud service providers. Cloud consumers can successfully reduce total cost of resource provisioning using Optimal Cloud Resource Provisioning (OCRP) algorithm in cloud computing environment. The two provisioning plans are reservation and on-demand, used for computing resources which is offered by cloud providers to cloud consumers. The cost of utilizing computing resources provisioned by reservation plan is cheaper than that provisioned by on-demand plan, since a cloud consumer has to pay to provider in advance. This project proposes that the OCRP algorithm associated with rule based resource manager technique is used to increase the scalability of cloud on-demand services by dynamic placement of virtual machines to reduce the cost and also endow with secure accessing of resources from data centers and parameters like virtualized platforms, data or service management are monitored in the cloud environment.

Index Terms - Cloud Computing, Resource Provisioning, Virtual Machine Placement, Resource Manager, Eucalyptus.

1 INTRODUCTION

Cloud computing model used for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction.
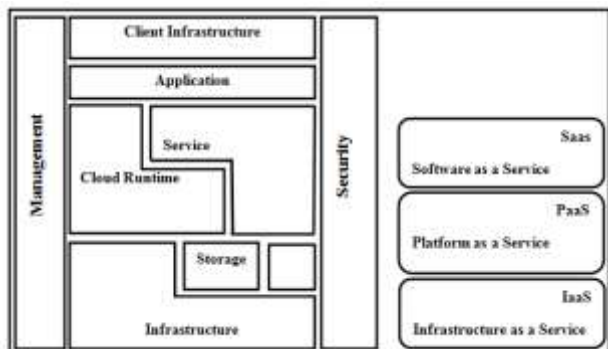

Figure 1: Cloud Computing Architecture

Figure 1 shows the cloud model which promotes the availability and is composed of five essential characteristics, three service models and four deployment models.

Cloud Computing infrastructure with the services are available "on-need" basis. The computing infrastructure includes hard disk, development platform, computing power, database or complete software applications. To access these resources from the cloud vendors, organizations do not need to make any large scale capital expenditures on their own. Organization need to pay only as much for the computing infrastructure as they use from the vendors whose hosting the services. The billing model of cloud computing is similar to the electricity payment that do on the basis of usage. The vendor is used for providing the cloud computing services and organization is used for users' of cloud computing services.

Virtualization is the creation of a virtual version of any of the resources like a server, an operating system, a storage device or network resources. It broadly describes the separation of a resource or request for a service from the physical entity.

Virtual machine placement is the process of mapping virtual machines to physical machines and it is the process of selecting the suitable host for the virtual machine. The process involves categorizing the resources requirements, virtual machines hardware, anticipated usage of resources and placement goal. The dynamic placement goal can either be maximizing the usage of available resources or it can be saving of power by being able to shut down some servers if the resources are not utilized at the computational time.

Infrastructure-as-a-Service (IaaS) is a computational service model provides virtual server instance to start, stop, access and configure their virtual servers and storage. In the enterprise, cloud computing allows a company to pay for only as much capacity as is needed for utilizing the services, and bring more online as soon as required because this pay-for-what-you-use model. It is a large-scale distributed computing paradigm in which a pool of computing resources is available to users (called cloud consumers) via the Internet [18]. Computing resources such as software, processing power, storage and network bandwidth are represented to cloud consumers as the accessible public utility services. In IaaS model, virtualization technologies can be used to provide resources to cloud consumers. They can specify the required software stack, e.g., operating systems and applications; then package them all together into Virtual Machines (VMs). The pricing in on-demand plan is charged by pay-per-use basis (e.g., 1 day) and on purchasing this on-demand plan, the consumers can dynamically provision resources at the moment when the resources are needed to fit the fluctuated and unpredictable demands. For reservation plan, pricing is charged by a onetime fee (e.g., 1 year) typically before the computing resource will be utilized by cloud consumer. With the reservation plan, the price to utilize resources is cheaper than that of the on-demand plan.

The Eucalyptus Cloud platform is open source software for building Amazon Web Services compatible private and hybrid clouds [17]. It pools together existing virtualized infrastructure to create cloud resources for network, compute and storage of services. It provides a highly robust and scalable Infrastructure as a Service solution. It is compatible with Amazon AWS - SOAP and REST interfaces and runs on multiple Linux distributions and supports Xen and KVM hypervisors. It has the ability to configure multiple clusters, each with private internal network addresses into a single Cloud.

Cloud Monitoring enables detect and triage performance problems, assign responsibility to the system or network engineering, application development group, automated tracking of the performance of the cloud services, automatic discovery of the deployed cloud topology, service availability, performance, utilization, throughput report, logs and alerts. Hyperic HQ is the Cloud monitoring platform for web application monitoring and performance management in physical, virtual and cloud environments. Hyperic monitors the health of Operating Systems, VMs, App Servers, Web Servers, Databases, etc.

The host level Virtual Machine hosting server security will be implemented by Meghdoot security which determines the Operating System (MAC) security Controls in BOSS environment. A host-based intrusion detection system (HIDS) monitors and analyzes the internals of a computing system as well as the network packets on its network interfaces.

In this paper the open source eucalyptus cloud environment will be utilized for accessing the resources to minimize the cost. With the reservation plan, the cloud consumers a priori reserve the resources in advance. As a result, the under provisioning problem can occur when the reserved sources are unable to fully meet the demand due to its uncertainty. Although this problem can be solved by provisioning more resources with on-demand plan to fit the extra demand, the high cost will be incurred due to more expensive price of resource provisioning with on-demand plan. In particular, an Optimal Cloud Resource Provisioning (OCRP) algorithm [1] and rule based resource manager technique [15] are used to minimize the total cost for provisioning resources in a certain time period. To make an optimal decision, the demand uncertainty from cloud consumer side and price uncertainty from cloud providers are taken into account to adjust the tradeoff between on-demand and oversubscribed costs. This optimal decision is obtained by means of a stochastic integer programming problem with multistage recourse. Benders decomposition [1] and sample-average approximation [1] are also used to solve the OCRP algorithm.

2 RELATED WORKS

The OCRP algorithm can provision computing resources for being used in multiple provisioning stages. The demand and price uncertainty is considered in OCRP and the different approaches to obtain the solution of the OCRP algorithm are considered including deterministic equivalent formulation, sample-average approximation, and Benders decomposition [1]. The resource provisioning problem is posed as one of sequential decision making under uncertainty and solved using a limited look ahead control scheme [8]. Cloud platforms host several independent applications on a shared resource pool with the ability to allocate computing power to applications on a per-demand basis. The use of server virtualization

techniques for such platforms provides great flexibility to consolidate several virtual machines on the same physical server, to resize a virtual machine capacity and to migrate virtual machine across physical servers. A key challenge for cloud providers is to automate the management of virtual servers while taking into account both high-level QoS requirements of hosted applications and resource management costs [11]. Virtualization technologies like VMware and Xen provide features to specify the minimum and maximum amount of resources that can be allocated to a virtual machine (VM) and a shares based mechanism for the hypervisor to distribute spare resources among contending VMs [12]. Dynamic consolidation is an approach that migrates tasks within a cluster as their computational requirements change, both to reduce the number of nodes that need to be active and to eliminate temporary overload situations. It allows Entropy, a consolidation manager for clusters to find mappings of tasks to nodes that are better than those found by heuristics based on local optimizations, and that are frequently globally optimal in the number of nodes. Because migration overhead is taken into account, Entropy chooses migrations that implemented efficiently by low performance overhead [13]. The objective [15] was the performance of Resource Manager on the basis of resource utilization and cost in hybrid cloud environment. Resource provisioning is important issue in cloud computing and in the environment of heterogeneous clouds.

The private cloud with confidentiality data configure according to users need. But the scalability of the private cloud limited. If the resources private clouds are busy in fulfilling other requests then new request cannot be fulfilled. The new requests are kept in waiting queue to process later. It take lot of delay to fulfill these requests and costly. The Rule Based Resource Manager is proposed for the Hybrid environment, which increases the scalability of private cloud on-demand and reduce the cost. Also set the time for public cloud and private cloud to fulfill the request and provide the services in time. In [16], an Optimal Virtual Machine Placement (OVMP) algorithm in which cloud providers can offer users two payment plans are reservation and on-demand plans for resource provisioning. Price of resources in reservation plan is generally cheaper than that in on-demand plan. However, since the reservation plan has to be acquired in advance, it may not fully meet future demands in which the on-demand plan can be used to guarantee the availability to the user. This algorithm can minimize the cost spending in each plan for hosting virtual machines in a multiple cloud provider environment under future demand and price uncertainty. OVMP algorithm makes a decision based on the optimal solution of Stochastic Integer Programming (SIP) to rent resources from cloud providers. This algorithm can be applied to provision resources in emerging cloud computing environments.

3 PROPOSED WORK

This proposed scheme has been analyzed on reducing together under provisioning and over provisioning complexity under the demand and cost vagueness in cloud computing environment and to motivate to explore a resource provisioning plan for cloud consumers. The Rule Based Resource Manager will be proven to be cost effective in money spent for using cloud resource. In demanding, an Optimal Cloud Resource Provisioning

(OCRP) algorithm is used to reduce the total charge for provisioning property in a certain time period and secure accessing of resources in the open source cloud computing environment.

The advantages of the proposed scheme are,

- In provisioning, the trouble can be solved by provisioning more resources with on-demand plan to fit the extra demand, the high cost will be incurred due to more expensive price of resource provisioning with on-demand plan.
- The cloud consumer to minimize the total cost of resource provisioning by reducing the on-demand cost and oversubscribed cost of under provisioning and over provisioning.
- Secure accessing of resources in the cloud.
- Dynamic placement of virtual machines for the fast response of the services.
- Monitoring the cloud computing environment.

If the reserved sources are unable to fully meet the demand due to its uncertainty, then the under provisioning problem can occur. Although this problem can be solved by provisioning more resources with on-demand plan to fit the extra demand, the high cost will be sustained due to more expensive price of resource provisioning with on-demand plan. On the other hand, the over provisioning problem can occur if the reserved resources are more than the actual demand in which part of a resource pool will be underutilized. It is important for the cloud consumer to minimize the total cost of resource provisioning by reducing the on-demand cost and oversubscribed cost of under provisioning and over provisioning. To achieve this goal, the optimal computing resource management is the critical issue in the cloud computing environment. It would be better to achieve this minimal provisioning cost with the rule based resource manager technique.
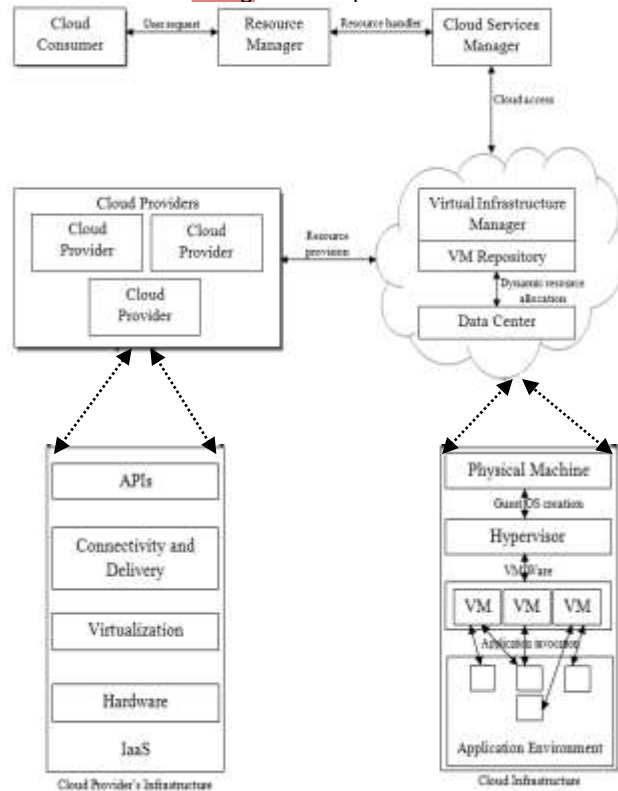

Figure 2: Dynamic resource provisioning by virtual machine placement architecture

Figure 2 depicts the interaction between the cloud consumer and the provider. The user request is handled by the resource manager and it is invoked on cloud service manager to handle the pool of resources by virtual machine repository to maintain different kind of resources resulting in dynamic resource allocation. The virtual machines is underlying by the physical machine which has a set of processes to be efficiently handled by the set of virtual machines in the application environment. The cloud provider provides with the IaaS model and features for providing the security for accessing the services.

This architecture consists of following modules:

- Configuring Eucalyptus Cloud
- Optimal Cloud Resource Provisioning (OCRP)
- Rule-Based Resource Manager Technique
- Dynamic Virtual Machine Placement
- Secure Resource Provisioning

3.1 Configuring Eucalyptus Cloud

In this paper, the eucalyptus configuration is done by the System which is booted in BOSS GNU/Linux 4.0 with Xen Kernel. Installation of packages in BOSS Linux requires root user access. Identify the Sources.list file should contain the necessary repository URL. The iproute and iptables packages should be installed in the Cloud Host. The Node Machine is required to have Xen Hypervisor installed with the packages of linux-system image, xen-utils-3.2-1 and xenstore-utils, xenwatch, xen-shell and xen-tools. While Installation of the Eucalyptus Node Packages the machine is required to be booted in Xen Kernel and Xen Management Tools must be working. Download the eucalyptus and source packages for executing the resources as the services.

The eucalyptus cloud environment can be held up on adding the number of nodes and on configuring the cloud host and cluster controller services by registering a node and accessing it through cloud portal.

3.2 Optimal Cloud Resource Provisioning (OCRP)

The cloud broker considers both reservation and on-demand plans for provisioning resources. These resources are used in different time intervals, also called provisioning phases. There are three provisioning phases: reservation, expending, and on-demand phases. A provisioning stage is the time epoch when the cloud broker makes a decision to provision resources by purchasing reservation and/or on-demand plans, and also allocates VMs to cloud providers for utilizing the provisioned resources. A cloud provider can offer the multiple reservation plans for consumer with different reservation contracts. Each reservation contract refers to the advance reservation of resources with the specific time duration of usage.

The general form of stochastic integer program of the OCRP algorithm is formulated in (1) and (2). The objective function (1) is to minimize the cloud consumer's total provisioning cost.

$$z = \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} c_{ijk}^{(R)} x_{ijk}^{(R)} + \mathbb{E}_{\Omega} \left[Q\left(x_{ijk}^{(R)}, \omega \right) \right].$$

$$(1)$$

subject to:

$$x_{ijk}^{(\mathrm{R})} \in \mathbb{N}_0, \quad \forall i \in \mathcal{I}, \forall j \in \mathcal{J}, \forall k \in \mathcal{K}.$$

(2)

Where,

I - Set of virtual machine (VM) classes while $i \in I$ denotes the VM class index.

J - set of cloud providers while $j \in J$ denotes the cloud provider index.

K - Set of reservation contracts while $k \in K$ denotes the reservation contract index.

Ω - Set of scenarios while $\omega \in \Omega$ denotes the scenario index.

$c_{ijk}^{(\mathrm{R})}$ - Reservation cost subscribed to reservation contract k charged by clod provider j to cloud consumer's VM class i in the reservation stage.

$x_{ijk}^{(\mathrm{R})}$ - Decision variable representing the number of VMs in class i provisioned in reservation phase subscribed to reservation contract k offered by cloud provider j in the reservation stage.

The Benders decomposition algorithm [1] is applied to solve the stochastic programming problem. The goal of this algorithm is to break down the optimization problem into multiple smaller problems which can be solved independently and parallelly. As a result, the time to obtain the solution of the OCRP algorithm can be reduced. The Benders decomposition algorithm can decompose integer programming problems with complicating variables into two major problems as master problem and subproblem.

The number of scenarios is numerous; it may not be efficient to obtain the solution of the OCRP algorithm by solving the stochastic programming formulation directly if all scenarios in the problem are considered. To address this complexity issue, the Sample-Average Approximation (SAA) approach is applied. This approach selects a set of scenarios, e.g., N scenarios, where N is smaller than the total number of scenarios $|\Omega|$. Then, these N scenarios can be solved in a deterministic equivalent formulation. The optimal solution can be obtained if N is large enough which has already been verified numerically[1].

3.3 Rule-Based Resource Manager Technique

Rule Based Resources Manager for successfully utilizing the private cloud resources and considering the security requirements of applications and data. With resource manager a private cloud can be scaled up to allocate resources on-demand even if private cloud overloaded. Also the scalability beyond the capacity of private cloud is achieved by using public cloud resources. The Resource Manager recognizes the suitable cloud to be used to fulfill a request.

The high priority request always access resources from the private cloud itself, because it have confidential information. Next low priority requests can be fulfilled from either public cloud or private cloud. But if the private cloud resources are available, it must be used first as these resources are possess by the enterprise and should be utilized. Sometimes high priority request fulfilled by private cloud but its resources are already assigned to fulfill previous requests of low and high priority. In this situation, the already allocated low priority requests and reallocate these low priority requests for which the remaining cost on public cloud is minimum to public cloud.

Steps to process the incoming requests:

- If new request is available, then redirects new virtual machine request to the private cloud.
- If the new incoming request has high priority availability, then check any lower priority process with lesser time to get completed is handling by the private server. If true, make the new incoming request to be waited in the waiting queue.
- If no space available in the specific virtual machine, redirects the process to free non processing virtual machine at the public cloud.

Rule-based Service Level Agreement (RBSLA) follows a knowledge based approach and uses RuleML to specify the SLA. The traditional way to support Service Level Agreement is Webservice SLA(WSLA) which is designed to capture service level agreements and comprises of mainly three entities such as parties (service provider, service consumer and third parties), SLA parameters(metrics measurement - resource and composite and functions) and Service Level Objectives (formal expressions). An established WSLA contains the following major sections:

- Parties - It comprises of two parties, they are supporting parties and signatory parties. Signatory parties are the service provider and the service consumer. Supporting parties are the third parties that come into picture when signatory parties decide to delegate certain tasks such as measuring SLA parameters.
- Service Definitions - A service dentition contains the description of the service providers interface. Services are represented by service objects. Each service object associates with one or more SLA parameters.
- Obligations - It contains the conditions and the action guarantees.

3.4 Dynamic Virtual Machine Placement

The virtual machine placement has been classified into power based and application QoS based approach. In which the stochastic integer programming will be processed by power based mechanism on dynamic migration of virtual machines. The Optimal Virtual Machine Placement (OVMP) algorithm provisioned the resources offered by multiple cloud providers. The algorithm is based on an IaaS model which leverages virtualization technologies. This OVMP algorithm minimizes the total cost of resource provision in a cloud computing environment. The tradeoff between the advance reservation of resources and the allocation of on-demand resources is adjusted to be optimal. Dynamic allocation is implemented on shorter timescales, preferably shorter than periods of significant variability of the resource demand. The Placement algorithms run in the background of the application processes collecting data.

3.5 Secure Resource Provisioning

The reasons for security in cloud are,

- Cloud Computing is Internet based computing.
- Common resources will be shared among many users.
- Data is moving from your data centre to cloud provider premises.
- It involves security of both physical and virtual resources.

- Physical resources security is similar to traditional security approaches.
- Virtual resource security needs special attention.

Security issues in both virtualization and cloud layers. The Virtualization will be achieved with the help of Hypervisor. Since Virtualization layer acts as a basement for cloud, many attackers are targeting it and host/platform security, security between guests, security between host/guests, data loss and management interfaces. In cloud layer, abuse and nefarious use of cloud computing, insecure application programming interfaces, malicious insiders, shared technology issues, data loss or leakage, account or service hijacking and unknown risk profile.

The meghdoot interface is used to handle the issues of security in cloud computing environment. and also the cloud will be monitored by the Hyperic HQ tool monitors the parameters like data centre management – hardware, network & devices, software management database and application, virtualized platforms, power and cooling factors, cloud services management, user tracking and security, real time performance - load, response time and throughput scalability and Operations - process, service availability, capacity and resource utilization.

4 CONCLUSION AND FUTURE WORK

The OCRP algorithm associated with rule based resource manager technique is used to increase the scalability of cloud on-demand services by dynamic virtual machines placement to minimize the cost. The optimal solution will be obtained from OCRP algorithm with rule based resource manager technique by formulating the stochastic integer programming with multistage recourse and also applied Benders decomposition approach to divide an OCRP problem into subproblems which can be solved parallelly. Furthermore, it has applied the Sample Average Approximation approach for solving the OCRP problem with a large set of scenarios.

Eucalyptus cloud is used for providing the resources offered by multiple cloud providers in which Hyperic HQ tool helps to monitor the transactions for accessing the services and Meghdoot for virtual machine security. It can effectively achieve an estimated optimal solution even the problem size is greatly large in number.

This work can be extended to incorporate the scenario reduction techniques to reduce the number of scenarios of uncertainty parameters to optimize the cost minimally with dynamic placement of virtual machines and also it will be implemented in more than one open source trusted cloud computing environment to calculate the response time and performance of the computational resources by cloud monitoring.

ACKNOWLEDGMENTS

REFERENCES

[1] Sivadon Chaisiri and Bu-Sung Lee, "Optimization of Resource Provisioning Cost in Cloud Computing," IEEE Transactions on Services Computing, VOL. 5, NO.2, APRIL-JUNE 2012.

[2] F.V. Louveaux, "Stochastic Integer Programming," Handbooks in OR & MS, vol. 10, pp. 213-266, 2003.

[3] A.J. Conejo, E. Castillo, and R. Garcia-Bertrand, "Linear Programming: Complicating Variables," Decomposition Techniques in Mathematical Programming, chapter 3, pp. 107-139, Springer, 2006.

[4] J. Linderoth, A. Shapiro, and S. Wright, "The Empirical Behavior of Sampling Methods for Stochastic Programming," Ann. Operational Research, vol. 142, no. 1, pp. 215-241, 2006.

[5] K. Miyashita, K. Masuda, and F. Higashitani, "Coordinating Service Allocation through Flexible Reservation," IEEE Trans. Services Computing, vol. 1, no. 2, pp. 117-128, Apr.-June 2008.

[6] G. Juve and E. Deelman, "Resource Provisioning Options for Large-Scale Scientific Workflows," Proc. IEEE Fourth Int'l Conf. e-Science, 2008.

[7] Y. Kee and C. Kesselman, "Grid Resource Abstraction, Virtualization, and Provisioning for Time-Target Applications," Proc. IEEE Int'l Symp. Cluster Computing and the Grid, 2008.

[8] D. Kusic and N. Kandasamy, "Risk-Aware Limited Lookahead Control for Dynamic Resource Provisioning in Enterprise Computing Systems," Proc. IEEE Int'l Conf. Autonomic Computing, 2006.

[9] J. Chen, G. Soundararajan, and C. Amza, "Autonomic Provisioning of Backend Databases in Dynamic Content Web Servers," Proc. IEEE Int'l Conf. Autonomic Computing, 2006.

[10] L. Grit, D. Irwin, A. Yumerefendi, and J. Chase, "Virtual Machine Hosting for Networked Clusters: Building the Foundations for Autonomic Orchestration," Proc. IEEE Int'l Workshop Virtualization Technology in Distributed Computing, 2006.

[11] H.N. Van, F.D. Tran, and J.-M. Menaud, "SLA-Aware Virtual Resource Management for Cloud Infrastructures," Proc. IEEE Ninth Int'l Conf. Computer and Information Technology, 2009.

[12] M. Cardosa, M.R. Korupolu, and A. Singh, "Shares and Utilities Based Power Consolidation in Virtualized Server Environments," Proc. IFIP/IEEE 11th Int'l Conf. Symp. Integrated Network Management (IM '09), 2009

[13] F. Hermenier, X. Lorca, and J.-M. Menaud, "Entropy: A Consolidation Manager for Clusters," Proc. ACM SIGPLAN/ SIGOPS Int'l Conf. Virtual Execution Environments (VEE '09), 2009.

[14] W.K. Mak, D.P. Morton, and R.K. Wood, "Monte Carlo Bounding Techniques for Determining Solution Quality in Stochastic Programs," Operations Research Letter, vol. 24, pp. 47-56, 1999.

[15] Rajkamal Kaur Grewal & Pushpendra Kumar Pateriya, "A Rule-based Approach for Effective Resource Provisioning in Hybrid Cloud Environment", International Journal of Computer Science and Informatics ISSN (PRINT): 2231 –5292, Vol-1, Iss-4, 2012.

[16] S. Chaisiri, B.S. Lee, and D. Niyato, "Optimal Virtual Machine Placement across Multiple Cloud Providers," Proc. IEEE Asia-Pacific Services Computing Conf. (APSCC), 2009.

[17] Eucalyptus, http://www.eucalyptus.com/eucalyptus-cloud

[18] I. Foster, Y. Zhao, and S. Lu, "Cloud Computing and Grid Computing 360-Degree Compared," Proc. Grid Computing Environments Workshop (GCE '08), 2008.

[19] G.B. Dantzig and G. Infangerm, "Large-Scale Stochastic Linear Programs: Importance Sampling and Benders Decomposition," Proc. IMACS World Congress on Computation and Applied Math., 1991.

[20] H. Heitsch and W. Romisch, "Scenario Reduction Algorithms in Stochastic Programming," J. Computational Optimization and Applications, vol. 24, pp. 187-206, 2003.

Aarthee S received the Master of Computer Applications degree from Anna University, Tiruchirappalli, TamilNadu in 2010. She is currently working toward the M.E. Computer Science and Engineering degree at Sudharsan Engineering College, Pudukkottai, TamilNadu, India. Her current research interests include cloud computing and data mining.

R.Venkatesan received the B.E. degree from Bharathiar University, Coimbatore, TamilNadu in 2004 and the M.E. degree in Computer Science and Engineering from the Anna University, Chennai, TamilNadu in 2007. He is currently an assistant professor in Sudharsan Engineering College, Pudukkottai, TamilNadu, India. His current research interests includes image processing, networks and cloud computing.