Optimization of Cost in Cloud Computing Using OCRP Algorithm

A.Poobalan[#1], V.Selvi[*2]

[#] Assistant Professor, Department of Computer Science and Engineering, University college of Engineering, Panruti, Tamil Nadu, India.

[*] PG Scholar, Department of Computer Science and Engineering, University College of Engineering, BIT Camps, Trichy, Tamil Nadu, India.

***Abstract*—In cloud computing, cloud providers can offer cloud consumers two provisioning plans for computing resources, namely reservation and on-demand plans. Reservation plan is cheaper than that provisioned by on-demand plan, since cloud consumer has to pay to provider in advance. By using reservation plan, the consumer can reduce the total resource provisioning cost. Even though, the best advance reservation of resources is difficult to be achieved due to uncertainty of consumer's future demand and providers' resource prices. Due to this problem, an optimal cloud resource provisioning (OCRP) algorithm is proposed by formulating a stochastic programming model. The OCRP algorithm can provision computing resources for being used in multiple provisioning stages as well as a long-term plan, e.g., four stages in a quarter plan and twelve stages in a yearly plan. The demand and price uncertainty is considered in OCRP. In this paper, different approaches are measured including deterministic equivalent formulation, sample-average approximation, and Benders decomposition in OCRP algorithm.**

***Keywords*— Virtualization, virtual machine placement, Deterministic equivalent formulation, Sample-average approximation, Benders decomposition.**

I. Introduction

Cloud computing is a large-scale distributed computing paradigm in which a pool of computing resources is available to cloud consumers via the Internet [1]. Computing resources, such as processing power, storage, software, and network bandwidth, are represented to cloud consumers as the available public utility services. Infrastructure-as-a-Service (IaaS) is a computational service model widely used in the cloud computing paradigm. Here in this method, virtualization technologies can be used to supply resources to cloud customers. The clients can specify the required software stack, e.g., operating systems and applications; then enclose them all together into virtual machines (VMs). The hardware requirement of VMs can also be adjusted by the consumers. At last, these VMs will be outsourced to host in computing environments.

With the reservation plan, the cloud consumers will previously reserve the resources in advance. There may occur underprovisioning problem when the reserved resources are unable to fully meet the demand due to its uncertainty. Although this problem can be solved by ordering extra resources by the help of 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. At the same time, the overprovisioning problem can occur when the reserved resources are more than the real demand in which part of a resource pool will be underutilized. The main part of the cloud consumer is to minimize the total cost of resource provisioning by reducing the on-demand cost and oversubscribed cost of underprovisioning and overprovisioning. To reach this goal, the optimal computing resource management is the critical issue. Both underprovisioning and overprovisioning problems must be minimized under the demand and price uncertainty in cloud computing environments is our motivation to explore a resource provisioning policy for cloud consumers. To overcome the problems such as underprovisioning and overprovisioning an algorithm is used called optimal cloud resource provisioning (OCRP) algorithm. It is proposed to minimize the total cost for provisioning resources in a particular time period. In order to the cost of the resource provisioning in cloud computing, the demand uncertainty from cloud consumer side and price uncertainty from cloud providers are considered to adjust the trade-off between on-demand plan and oversubscribed costs.

II. RELATED WORK

The optimal virtual machine placement (OVMP) algorithm [2] was used before OCRP algorithm. This OVMP algorithm was used to solve the problem of both resource provisioning for cloud consumers and VM placement in two provisioning stages. To improve the number of provisioning stages of the previous work, we introduce the algorithm called OCRP algorithm in this paper. This OCRP algorithm improves the number of provision stages as well as reduces the cost compared to previous work. First, the problem is generalized into the multiple stage formulation. Second, the different approaches to obtain the solution of computing resource provisioning are considered. Finally, the performance evaluation is extended to consider various realistic scenarios.

III. SYSTEM MODEL AND ASSUMPTION

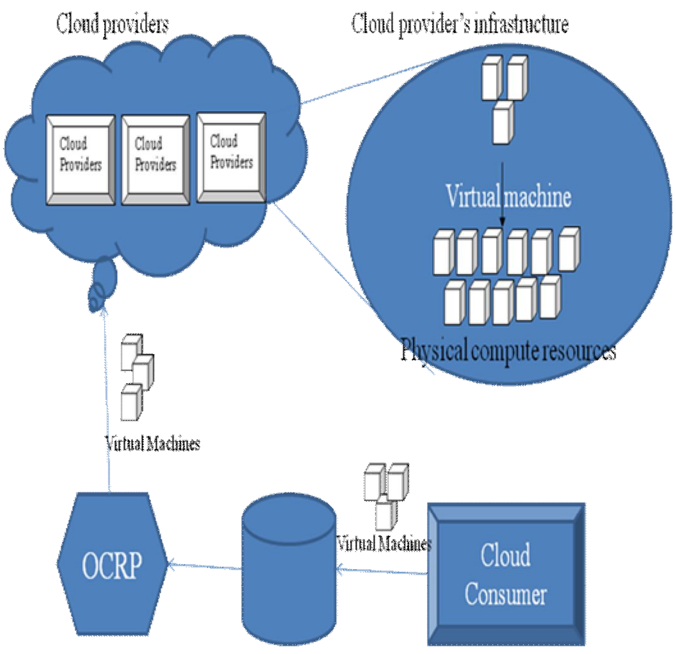

Fig. 1 System model of OCRP Algorithm

The system model and assumption of cloud computing environment consists of four main components, namely cloud consumer, virtual machine (VM) repository, cloud providers, and cloud broker as shown in the Fig.1. The cloud consumer can order the requirements to execute jobs from cloud providers. Computing resources has to be provisioned from cloud providers, before the jobs ordered by cloud consumers are executed. To get such resources, first the consumer creates VMs integrated with software required by the jobs. VM repository stores the created VMs. Then, the VMs can be hosted on cloud providers' infrastructures. From the cloud provider infrastructures the VMs can utilize the resources. In Fig. 1, the cloud broker is located in the cloud consumer's site. Thus cloud broker is responsible on behalf of the cloud consumer for provision resources for hosting the Virtual Machines. And also, the broker can allocate the VMs originally stored in the VM repository to suitable cloud providers. Here the broker uses the OCRP algorithm to reduce the cost of resource provisioning in cloud computing environment.

IV. STOCHASTIC PROGRAMMING MODEL

In this stochastic programming model section, the stochastic programming with multistage recourse [3] is presented as the important part of the OCRP algorithm. The original form of stochastic integer programming formulation is derived first in OCRP algorithm. Then, the formulation is transformed into the deterministic equivalent formulation (DEF). This deterministic equivalent formulation can be solved by traditional optimization solver software.

V. BENDERS DECOMPOSITION

The Benders decomposition algorithm [4] is applied in this part to solve the stochastic programming problem formulated in Section IV. The main aim of this algorithm is to split down the optimization problem into number smaller problems which can be solved separately and simultaneously. Because of this benders decomposition algorithm the time can be reduced to obtain the solution of the OCRP algorithm. The Benders decomposition algorithm can decompose integer programming problems with complicating variables into two major problems: one is master problem and another one is sub problem. Flow chart for benders decomposition algorithm was shown in Fig.2.

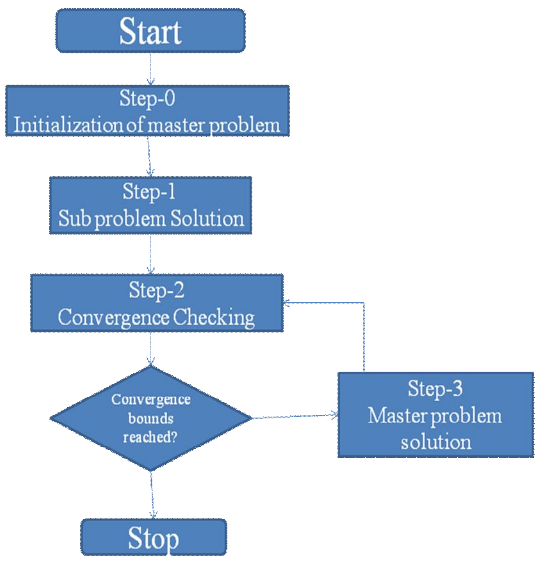

Fig. 2 Flow chart for Benders decomposition algorithm

VI. SAMPLE-AVERAGE APPROXIMATION

When the number of scenarios is numerous, it may be difficult to obtain the solution of the OCRP algorithm by solving the stochastic programming formulation defined in [5]-[10], [11], [2] directly if all scenarios in the problem are taken into account. In order to avoid this major problem, the sample-average approximation (SAA) approach is applied [12]. In this approach, it selects a set of scenarios, e.g., N scenarios, where N is smaller than the total number of scenarios |Ω|. Then, deterministic equivalent formulation can be used to solve these N scenarios. The optimal solution can be obtained if N is large enough which can be verified numerically.

VII. CONCLUSION

In this paper, we have proposed an optimal cloud resource provisioning (OCRP) algorithm to provision resources offered by multiple cloud providers. Stochastic integer programming with multistage recourse was solved and formulated to get the optimal solution of OCRP algorithm. And Benders decomposition approach was also applied to divide an OCRP problem into many sub problems which can be solved simultaneously to save the time. And also, we have applied the Sample Average Approximation approach for solving the OCRP problem with a large set of scenarios. The SAA approach can successfully achieve an expected optimal solution even though the problem size is very much large. The performance evaluation of the OCRP algorithm has been performed by numerical studies and simulations. According to the results, the algorithm can optimally correct the trade-off between reservation of resources and allocation of on-demand resources by cloud consumers. The OCRP algorithm can effectively save the cost of resource provisioning for cloud consumers. It can be used as resource provisioning tool in cloud computing environment which reduces the cost effectively.

REFERENCES

[1] I. Foster, Y. Zhao, and S. Lu, “Cloud Computing and Grid Computing 360-Degree Compared,” Proc. Grid Computing Environments Workshop (GCE ’08), 2008.
[2] 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.
[3] F.V. Louveaux, “Stochastic Integer Programming,” Handbooks in OR & MS, vol. 10, pp. 213-266, 2003.
[4] A.J. Conejo, E. Castillo, and R. Garcı´a-Bertrand, “Linear Programming: Complicating Variables,” Decomposition Techniques in Mathematical Programming, chapter 3, pp. 107-139, Springer, 2006.
[5] 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.
[6] 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.
[7] 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.
[8] 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.
[9] N. Bobroff, A. Kochut, and K. Beaty, “Dynamic Placement of Virtual Machines for Managing SLA Violations,” Proc. IFIP/IEEE Int’l Symp. Integrated Network Management (IM ’07), pp. 119-128, May 2007.
[10] P. Jirutitijaroen and C. Singh, “Reliability Constrained Multi-Area Adequacy Planning Using Stochastic Programming with Sample-Average Approximations,” IEEE Trans. Power Systems, vol. 23, no. 2, pp. 504-513, May 2008.
[11] GNU Linear Programming Kit (GLPK), http://www.gnu.org/ software/glpk, 2012.
[12] 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.