Asia Pacific Journal of Research
A peer reviewed international Journal
Print -ISSN -2320 - 5504
Online-E-ISSN - 2347 - 4793

COST OPTIMIZATION IN DYNAMIC RESOURCE ALLOCATION USING VIRTUAL MACHINES FOR CLOUD COMPUTING ENVIRONMENT

S.Umamageswari [#1] ***M.C.Babu*** [*2]

[#]*PG Scholar, Department of Computer Science and Engineering*

St Peter's University, Avadi, Tamil Nadu, India

[*]*Assistant Professor , Department of Computer Science and Engineering*

St Peter's University,Avadi,Tamil Nadu,India

ABSTRACT

Cloud Computing is an emerging concept combining many fields of computing. It provides services, software and processing capacity over the internet. Deploying Virtual Machines in bulk is a very tedious and cumbersome job. Deploying a Virtual Machines on a Data Center involves various tasks such as installing the operating prepping the machine for a system ,adding a system to the domain and giving proper access to user. Cloud computing consists of two provisioning plan for allocating resources in cloud. They are Reservation plan and On-demand plan. Reservation plan is long term plan and On-demand plan is a short term plan. In On-demand plan the consumers can access resources at the time when they need. In Reservation plan the resources could be reserved earlier. Hence the cloud providers could charge the resources before consumers could use it. For on-demand pricing is done as pay-per-use basis but in reservation plan pricing is charged by one-time fee. With Reservation plan consumers could utilize the computing resources in a much cheaper amount than on-demand plan. The proposed scheme has two important features. (i) Virtual Machine Deployment Algorithm based on the Hungarian Algorithm to support a concurrent Deployment with multiple Virtual Machine instance and monitoring the Resources (ii) Implementation of Optimal Cloud Resources Provisioning to make an optimal decision of Resource Provisioning in the Cloud Environment. Experimental results show the significant improvement in load balancing and reduce the reservation cost and expending cost.

Index Terms- *Hungarian Algorithm, Stochastic Integer Programming (SIP), Determinist Equivalent Formulation (DEF), Sample Average Approximation*

1. INTRODUCTION

Cloud Computing become a de facto standard for computing, infrastructure as a services (IaaS) has been emerged as an important paradigm in IT area. By applying this paradigm we can abstract the underlying physical resource such a CPUs, Memories and Storage and offer these Virtual Resource to users in the formal Virtual Machine. Multiple Virtual Machine are able to run on a single physical machine. Multiple VMs are able to run on a Single Physical Machine (PM). Another important issues in Cloud computing is provisioning method for allocating resources to cloud consumers. Cloud computing consists of two provisioning plan for allocating resources in cloud. They are Reservation plan and On-demand plan. Reservation plan is long term plan and On-demand plan is a short term plan. In On-demand plan the consumers can access resources at the time when they need. In Reservation plan the resources could be reserved earlier. Hence the cloud providers could charge the resources before consumers could use it. In on-demand pricing is done as pay-per-use basis but in reservation plan pricing is charged by one-time fee. With Reservation plan consumers could utilize the computing resources in a much cheaper amount than on-demand plan. Even though with the reservation plan the cloud consumer could use the resources in advance some problems could occur with it. One is the under provisioning problem in which the consumers could not fully meet the required resources due to uncertainty of allocating resources. Other problem with reservation plan is over provisioning of resources, where the reserved resources will be more than what actually needed. Hence the resources reserved will not be fully used. The goal is to achieve an optimal solution for provisioning resource which is the most critical part in cloud computing. To make an optimal decision, the demand, price, idle-time and waiting-time uncertainties are taken into account to adjust the tradeoffs between on-demand and oversubscribed costs. The Bender's Decomposition is applied to divide the resource optimization problem into many sub problems to decrease the on demand cost and Reservation Cost. Scenario Reduction Technique are applied to reduce problem by reducing number of Scenarios. This will decrease Reservation Price and Expending Price.

2. RELATED WORKS

In [1] "On networking and computing environments' integration: A novel mobile cloud resources provisioning approach" by Skoutas, D.N. , Skianis, C. A Mobile Cloud Resources Provisioning (MCRP) scheme, which is flexible enough to adapt to the various general MCC reference use cases being described. The main novelty feature of the employed MCC Service Admission Control algorithm lies in the fact that it jointly handles radio and computing resources rather than confronting the problem as two independent resource management sub-problems In [2] "Optimal resource allocation for multimedia cloud in priority service scheme "by He, Yifeng; Guan, Ling employ the queuing model to optimize the resource allocation for multimedia cloud in priority service scheme**.** Specifically, formulate and solve the resource cost minimization problem and the service response time minimization problem respectively. In [3] "Optimization of Resource Provisioning Cost in Cloud Computing "by Bu-Sung Lee **,**Niyato, D. 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 to obtain the solution of the OCRP algorithm are considered including deterministic

ISSN: 2320-5504, E-ISSN-2347-4793

equivalent formulation, sample-average approximation, and Benders decomposition. Numerical studies are extensively performed in which the results clearly show that with the OCRP algorithm, cloud consumer can successfully minimize total cost of resource provisioning in cloud computing environments. The OCRP algorithm was proposed in [4].The OCRP algorithm can find an optimal solution for resource provisioning and VM placement. It uses only two uncertainties only viz., demand and price. Here in this paper RCRP algorithm is used which is an extension of OCRP where four uncertainty factors are considered. Grid provides services which are not of desired quality. One of the major drawbacks of grid is single point of failure where one unit on the grid degrades which will cause the entire system to degrade. Hence [5] suggests cloud which is used for adaption of various services. The benefit of cloud is that it will avoid single point of failure and also will decrease hardware cost

3. ARCHITECTURE OF IaaS FRONT END-NODE

The VM deployment process maps VMs to PMs. In this process a PM's computing resource and required VMs resource are two major inputs of the placement problem. A wrong VM deployment makes inefficient use of computing resources which causes low resource utilization and imbalanced job loading. Thus the performance of the VM deployment process is important factor of efficient and load balancing performance of cloud service

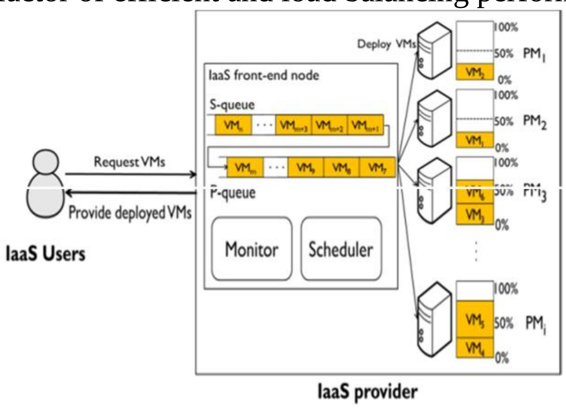

Fig 1. Architecture of IasS Front End Node

In this architecture a secondary job queue in addition to a primary job queue for incoming job requests. Hungarian algorithm is applied for distribution of job in the primary queue .While the job in the primary queue is in distribution, the secondary queue is holding incoming VM request from users. The size of the secondary queue is static and the one for the primary queue is dynamic. The size of primary queue should be less than the number of PMs that have enough resources for VM requests, the primary queue cannot store the large number of incoming VM requests than the number of PMs. That lead to a necessity of having a secondary queue to keep the incoming VM requests. The front end node sends the incoming VM request in the secondary queue. VM requests are transferred to the primary queue if it is empty. Then the front end node distributes VM requests stored in the primary queue to PMs by using the proposed adaptive VM deployment algorithm. In the proposed VM deployment algorithm we considered the VM execution time and two computing resources a CPU and memory to calculate the core

ISSN: 2320-5504, E-ISSN-2347-4793

3.1 HUNGARIAN METHOD

The Hungarian method is a combinatorial optimization algorithm that solves the assignment problem in polynomial time and which anticipated later primal-dual methods. Given VM and tasks, and an $n \times n$ matrix containing the cost of assigning each VM to a task, find the cost minimizing assignment. First the problem is written in the form of a matrix as given below

a1	a2	a3	a4
b1	b2	b3	b4
c1	c2	c3	c4
d1	d2	d3	d4

Where a, b, c and d are the VM who have to perform tasks 1, 2, 3 and 4. a1, a2, a3, a4 denote the penalties incurred when VM "a" does task 1, 2, 3, 4 respectively. The same holds true for the other symbols as well. The matrix is square, so each VM can perform only one task. Then we perform row operations on the matrix. To do this, the lowest of all *ai* (i belonging to 1-4) is taken and is subtracted from each element in that row. This will lead to at least one zero in that row (We get multiple zeros when there are two equal elements which also happen to be the lowest in that row). This procedure is repeated for all rows. We now have a matrix with at least one zero per row. Now we try to assign tasks to VM such that each VM is doing only one task and the penalty incurred in each case is zero.

0	A2'	0'	A4'
b1'	B2'	B3'	0'
0'	C2'	C3'	C4'
d1'	0'	D3'	D4

ISSN: 2320-5504, E-ISSN-2347-4793

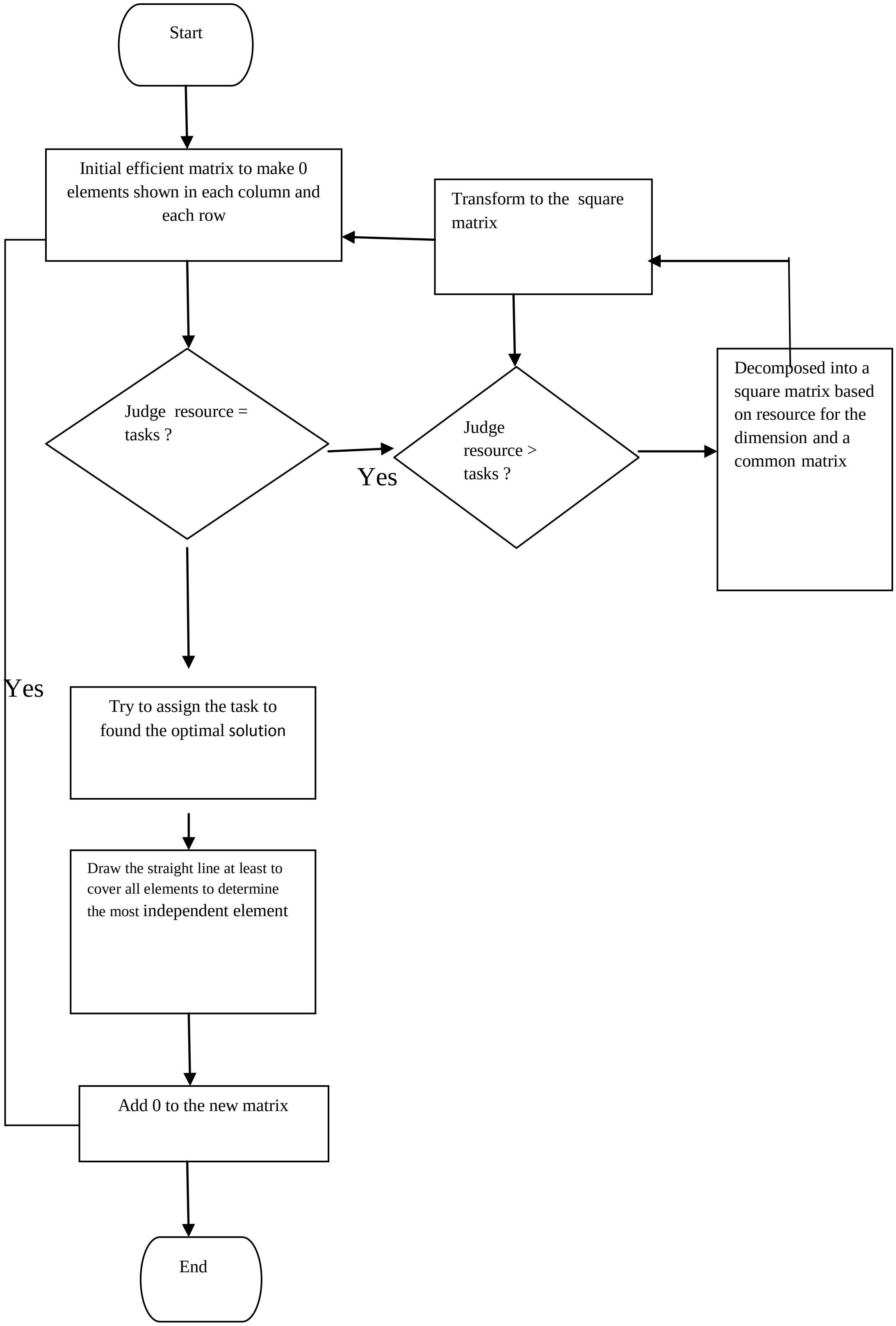

Fig 2. Steps in Hungarian Algorithm

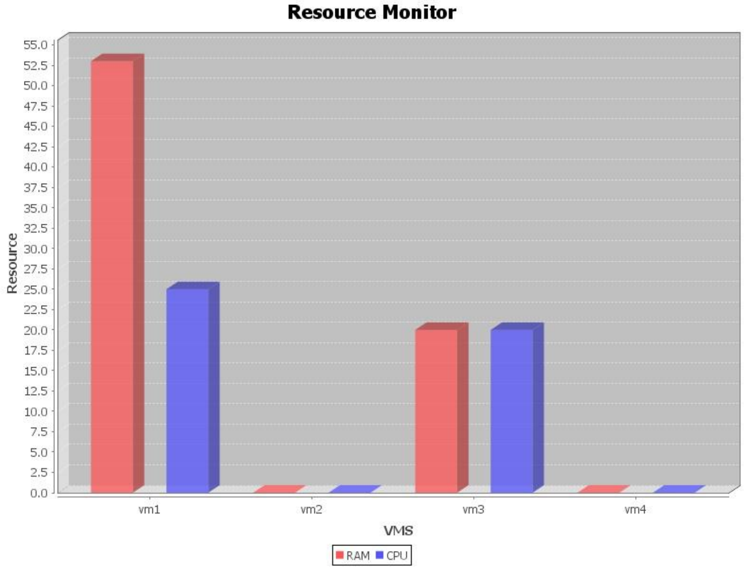

Fig 3. Resource Monitoring

4. OPTIMAL CLOUD RESOURC PROVISIONING

In Cloud computing environment the cost optimization problem draws significant of optimizing resource price and how to optimally provision cloud resources to meet service requirements. In Cloud Environment On Demand Cost, Reservation Cost and Expending Cost are the major areas to be used for finding the Optimal resource Cost. The stochastic programming (SPI) is used for finding the optimal resource cost under uncertainty. The Deterministic Equivalent Formulation (DEF) algorithm is used for solving linear mathematical optimization programming script errors is used to reduce the Cost in the On Demand. The Benders Decomposition algorithm is used for break down the optimization problems which they are reduced to many sub problems. It is used to reduce the on demand cost and reservation cost during the resource provisioning stage. The Sample Average Approximation (SAA) Algorithm can reduce the problem scenarios to obtain optimal resource provisioning cost. It is used to reduce the reservation cost and expending cost

4.1 PROPOSED SCHEME

The overall process of the proposed system is depicted in Fig. 4 which contains resource provisioning model, stochastic integer programming, Deterministic Equivalent Formulation, Benders Decomposition, Sample Average Approximation.

ISSN: 2320-5504, E-ISSN-2347-4793

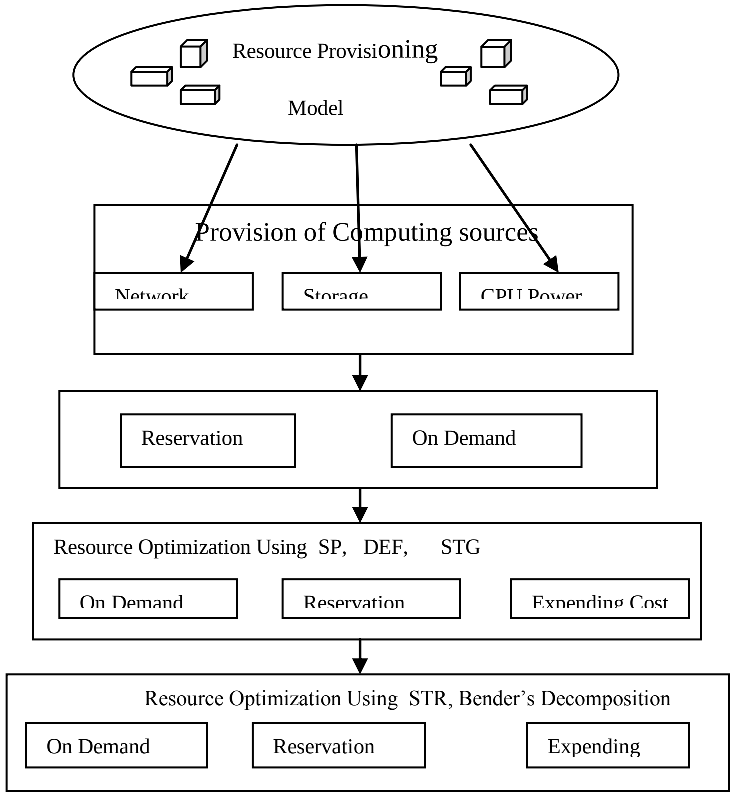

Fig 4. Overall Process of the Proposed System Model

4.1.1 RESOURCE PROVISIONING MODEL IN CLOUD

In this proposed system the VM ware Player is used to mount Linux Operating System in order to access the Open stack private cloud.. The computing resources are provisioned by using the resource provisioning model and the provision resources are network, storage, CPU processing power**.** The amount of resource types can be computing power in unit of CPU-hours, storage in unit of GBs/month, and network bandwidth for Internet data transfer in unit of GBs/month. In Virtual Machine repository each Virtual Machine class specifies the amount of resources in each resource type.

4.1.2 Key Notations of Resource Provisioning Model

I=Set of Virtual Machine (VM) classes while i I denotes the VM class index

J=Set of Cloud providers while jJ denotes the cloud provider index

K=Set of Reservation Contracts while k K denotes the reservation contract index

T=Set of provisioning stages while t Ɛ T denotes the provisioning stage index

R=Set of Resource types while r R denotes the resource type index

Ω=Set of Scenarios While ωƐΩ denotes the scenario index

Cijk(r)(ω) = Reservation Cost subscribed to reservation contract k charged by cloud provider j to cloud consumers VM class I in the first provisioning stage

Cijk(r)(ω) = Reservation Cost subscribed to reservation contract k charged by cloud provider j to cloud consumers VM class I in the first provisioning stage t and scenario ω

$c_{ijkt}^{(e)}(\omega)$ = Expending cost subscribed to reservation contract k charged by cloud provider j to cloud consumers VM class I in the first provisioning stage t and scenario ω

b_{ir} = Amount of resource type r required by VM class i

$d_{it(\omega)}$ = Number of Virtual Machines (VM) required to execute class i in provisioning stage t and scenario ω

$a_{jrt(\omega)}$ = Maximum capacity of resource type r that cloud provider j can offer to cloud consumer in provisioning stage t and scenario ω

Xijk(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 first provisioning stage

Xijk(R)(ω) = 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 provisioning stage t and scenario ω

Xijk(R)ω =Decision variable representing the number of VMs in class i run expending phase subscribed to Reservation contract k offered by cloud provider j in the provisioning stage t and scenario ω

4.1.3 Stochastic Integer Programming

Stochastic Programming is a Mathematical Programming about deciding under uncertainty. The goal of the stochastic programming model function is to minimize the cloud consumer's total resource provisioning cost. The deterministic optimization problems are developed with known parameters within certain bounds

$Minimize$: $\sum_{i\in i}\sum_{j\in j}\sum_{k\in k} c_{ijk}^{(R)} x_{ijk}^{(R)}$ + E_{Ω} [Q ($x_{ijk}^{(R)}$,ω)] **(1)**

Subject to: $x_{ijk}^{(R)} \in N_o$, Ɐi∈I,Ɐj∈J, Ɐk∈K **(2)**

Equation [1] is used to minimize cost in On Demand cost , Reservation cost and Expected cost. Two integer resource is used to solve the complexity of resource cost optimizations in uncertainty. Equation[2] is used to indicate the number of cloud provider, number of virtual machines and set of provisioning stages. In initial stage variables have to be determined before the realization of indefinite parameters In stage two indefinite values are represented as corrective measures. In stage one indefiniteness realization and functional level in stage two are the major issues. The aim is to choose the first-stage indefinite values in such a way that the sum of the stage first costs and the expected value of the random stage second costs is decreased.

4.1.4 Algorithm for Stochastic Integer Programming

N denotes the number of cloud providers

Step1: first get the sample reservation cost from cloud provider

Step2: initialize the cost for basic variables like Cijk(r) (ω)

for i = 1,2,... ,N do

for j = 1,2,... ,N do

for k = 1,2,... ,N do

z= scenario of reservation phase* decision variable of reservation phase I

(or)

z=Cijk(r)(ω)* Xijk(r)(ω)

I = min (decision variable of reservation phase, on demand phase, Expending phase) *c(y)
(or)
I=min (Xijk(r)(ω) * Xijk(o)(ω)* Xijk(e)(ω)
End for;
End for;
End for;
for i = 1,2,... ,N do
for j = 1,2,... ,N do
for k = 1,2,... ,N do
for t = 1,2,... ,N do
c(y)= scenario of reservation phase* decision variable of reservation phase + scenario of on demand phase* decision variable of on demand phase + scenario of expanding phase* decision variable of expanding phase (or)
c(y)= Cijkt(r)(ω)* Xijkt(r)(ω) + Cijkt(o)(ω)* Xijkt(o)(ω) + Cijkt(e)(ω)* Xijkt(e)(ω)
here, some constraints to be followed
Xijk(e)(ω)<=Xijk(r)(ω);
Xijk(R)= Xijkt(r)(ω);
Xijk(e)(ω)+Xijk(o)(ω) => dit(ω)
bir(Xijk(e)(ω)+Xijk(o)(ω)) <= ajrt(ω)
)
End for;
End for;
End for;

4.1.5 Deterministic Equivalent Formulation

Optimal First Stage Decision can be computed using deterministic equivalent formulation .The probability distributions of both price and demand can be used in deterministic equivalent formulation. In this optimization problem on demand cost is considered to be obtaining optimal solution of resource provisioning. The stochastic programming model uncertainty problems are solved here using deterministic formulation. In this formulation number of cloud provider are considered to optimize the on demand cost of resource provisioning.

4.1.6 Algorithm for Deterministic Equivalent Formulation

P (ω) - Probability distributions of both price and demand,
N denotes the number of cloud providers
Step 1: first get the sample reservation cost from cloud provider
Step 2: initialize the cost for basic variables like C ijk (r)(ω)
for i = 1,2,... ,N do
for j = 1,2,... ,N do
for k = 1,2,... ,N do
for t = 1,2,... ,N do
z=Cijk(R) (ω)* Xijk(R)(ω) + +(p(ω) * Cijkt(r)(ω)* Xijkt(r)(ω)) +p(ω) *(* Cijkt(o)(ω)* Xijkt(o)(ω) + Cijkt(e)(ω)* Xijkt(e)(ω));
where, Xijk(e)(ω)<=Xijk(r)(ω);
Xijk(R)= Xijkt(r)(ω);

ISSN: 2320-5504, E-ISSN-2347-4793

Xijk(e)(ω)+Xijk(o)(ω);=> dit(ω)
End for;
End for;
End for;
End for;

4.1.7 Benders Decomposition

The Benders decomposition algorithm can decompose integer programming problems with complicating variables into two major problems[1] master problem [2] sub problem. The master problems are constituted by the complicating variables and the sub problems are constituted by the other decision variables are solved, then lower and upper bounds are calculated by this approach.

4.1.8 Algorithm for Bender Decomposition

Step 1: split the problem to master and sub problem up to possibility
Step 2: Initialization of master problem,
Step 3: Solve the sub problem
for i = 1,2,... ,N do
for j = 1,2,... ,N do
for k = 1,2,... ,N do
S1= Zv(r) =Σ Σ Σ Cijk(r) + Σ p(ω) * Cijkt(r)(ω)* Xijkt(r)(ω)
Here, Xijkt(r)(ω) = Xijkt(fix)(ω) // it is for minimize the reservation cost
S2: Zv(o) =Σ Σ Σp(ω) *Cijk(o) * Xijk(o)(ω)
Here, Xijkt(o)(ω) = Xijkt(fix)(ω) // it is for minimize the on demand cost Where,
Xijk(e)(w) => dit(ω)
bir (Xijk(e)(w)) <= ajrt(ω)
End for;
End for;
End for;
Setp 4: Check the convergence condition zv'(ub)= zv *(e) – α v + zv *(r) +Σ zv *(o) (ω)
Step 5: If zv'(ub) then Stop the process (got optimal solution)
Else if;

Master Problem

Av =Σ Σ Σ Σ Σ((Yijktv(r)(ω)+ Yijktv(o)(ω)) * (Xijktv(e)(ω) - Xijktv'(e)(ω))

Here v'=1…v-1 and iteration counter be increased by v=v+ 1 for solve the sub problem and combine the master problem. After solving this master problem, Step-3 is repeated and the same iterative process continues

4.1.9 Sample Average Approximation(SAA)

In Proposed system we can achieve an optimal solution even the problem size is large using Sample Average Approximation. The estimation of Sample Average Approximations lower and upper bounds can yield tolerable solutions while the problems can be practically solved in timely manner. In the proposed system we can obtain optimal resource provisioning cost by Sample Average Approximation Approach

ISSN: 2320-5504, E-ISSN-2347-4793

4.1.10 Algorithm for SAA Upper & Lower Bound Estimation

N denotes number of scenarios N is smaller than the total number of scenarios $|\Omega|$

Step 1: Selects a set of scenarios -N

N scenarios can be solved in a deterministic equivalent formulation. The optimal solution can be obtained if N is large enough which can be verified numerically. The SAA approach is applied to approximate the expected cost in every considered provisioning stage.

Step 2: Problem is

$Zv(e) = \Sigma\ \Sigma\ \Sigma\ Cijk(R)(\omega) * Xijk(R)(\omega) + 1/N\ \Sigma\ \Sigma\ \Sigma\ \Sigma\ \Sigma\ Cijkt(r)(\ ) * Xijktv(r)(\ ) + 1/N\ \Sigma\ \Sigma\ \Sigma\ \Sigma\ (Cijkt(e)(\ ) * Xijktv(e)(\ ) + Cijkt(o)\ (\ ) * Xijktv(o)(\ )\ )$ It should be transformed into a deterministic equivalent formulation.

Step 3: Z* and x* denote the optimal objective function value and optimal solution of the original formulation. Z**N and x**N denote the optimal objective function value and optimal solution of the AP Formulation, where $Z^* \leq Z^{**}N$ both SAA upper and lower bounds on Z* with a certain confidence interval.

4.1.11 Scenario Tree Generation

One of the principal challenges in the field of stochastic programming handles with finding effective ways to assess the significance of atoms, and to make use of that data to reduce the tree of scenario's in such a way that the solution to the smaller best possible solution difficulty is not much dissimilar than the difficulty stated with the original tree. The Generation of Scenario Tree algorithm is a finite element technique that deals with this difficulty for the class of LSMP with random variables

4.1.12 Scenario Tree Reduction

The Scenario Tree Reduction (STR) algorithm is a finite element technique that deals with this problem for the class LSMP with random variables.

5. IMPLEMENTATION DETAILS

The performance of the proposed resource optimization framework is implemented using eclipse based java platform. Both Under provisioning and Over provisioning is solved by using different optimization problems under uncertainty. Resource Utilization and Cost Calculation is implemented in the Proposed system. In Fig 5. It displays the Customer details, resources utilization and cost calculation.

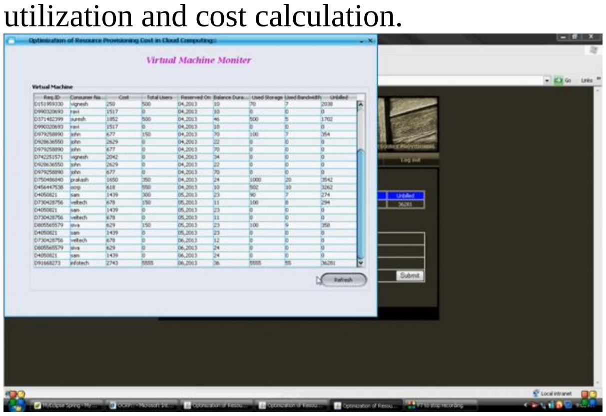

Fig 5. Optimization of Resource Provisioning Cost Computing

6. CONCLUSIONS

The open stack private cloud environment is configured by using oracle VM virtual box. Using Hungarian Method Virtual Machine Deployment implemented and Resources utilization is monitor in an effective manner The cloud resources are provisioned by using open stack resource provisioning model. The Two Stage Stochastic Integer Programming with recourse is applied to solve the complexity of optimization problems under uncertainty .Deterministic Equivalent Formulation (DEF) is used to solve the probability distribution of all scenarios to reduce the on demand cost. The Benders Decomposition is applied for break down the resource optimization problem into multiple sub problems. It is used to reduce the on demand cost and reservation cost during the resource provisioning stage. The Sample Average Approximation (SAA) is applied for reduce the problem scenarios in a resource optimization problem. This algorithm is used to reduce the reservation cost and expending cost.

7. REFERENCES

[1] Skoutas, D.N.; Skianis, C. "*On networking and computing environments' integration: A novel mobile cloud resources provisioning approach* "

[2] He, Yifeng; Guan, Ling "*Optimal resource allocation for multimedia cloud in priority service scheme* " in Circuits and Systems (ISCAS), 2012 IEEE International Symposium on 20-23 May 2012

[3] Bu-Sung Lee; Niyato, D. "*Optimization of Resource Provisioning Cost in Cloud Computing* "Services Computing, IEEE Transactions on April-June 2012 Volume: 5, Issue: 2 [5] "

[4] S. Chaisiri , B.S. Lee and D. Niyato, "Optimization of Resource Provisioning Cost in Cloud IEEE Transactions On Services Computing", Vol. 5, No. 2, April-June 2012.

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

[6]. " An optimal Solution of Resource Provisioning Cost in Cloud Computing Environments " by A Anto, Department of Computer Science and Engineering , Sri Krishna College of Technology, Coimbatore. On International Journal of computer Applications & information Technology.

[7]. Hongjae Kim,Munyoung Kang,Sanggil Kang, Sangyoon Oh " A Novel Adapative Virtual Machine Deployment Algorithm for Cloud Computing",Department of Computer Engineering, Ajou University, Suwon,South Korea

[8]. Qiang Li, "Applying Stochastic Integer Programming to Optimization of Resource Scheduling in Cloud Computing." Journal of Networks, Jul2012.