COST EFFECTUAL RESOURCE PROVISIONING USING IOVMA AND SUPERLATIVE INTEGER PROGRAMMING

P.Durgadevi[1], T.Chindrella Priyadharshini[2], M.Swarnalatha[3]

[1]Research Scholar, Anna university, Tamilnadu, (India)

[2,3]Assistant Prof, IT Dept, RMKCET, Puduvoyal, Tamilnadu,(India)

ABSTRACT

Cloud computing is networking technology in which set of resources are provided as services. Resource allocation in cloud computing is done based on client registration using services present in cloud computing. In resources allocation there is lot of query patterns for each client for acquiring the resources i.e. memory Consumption, CPU exploitation, and other resources are usage capabilities in cloud computing. For resource provisioning in cloud two mechanisms are used. 1) Reservation and 2) on-demand plan services. In terms of cost estimation process in the cloud there is challenging task in optimization of capacity utilization in deploying virtual machine placement assignment. In this paper, we proposed an Enhanced Optimal Virtual Machine Assignment Algorithm to implement resource provisioning operations. The proposed IOVMA Algorithm makes a decision process on cloud service provider and superlative integer programming to allocate resources from cloud service providers. These service professional accepts the cloud computing services with resource provisioning with suitable services. Our experimental results shows resources are provided with minimum cost constraints in emerging cloud computing environments.

Keywords: Networking, Superlative Integer Programming, Virtual Machine, On-Demand, Reservation.

I INTRODUCTION

Cloud Computing is a general term used to describe a new class of network based computing that takes place over the Internet, basically a step on from Utility Computing [1]. The word service in cloud computing is the concept of being able to use reusable, components across a third party server. Without using hardware and software the third party servers providing services according to the client requirement.

A number of characteristics define cloud data, applications services and infrastructure:

a. Remotely hosted: Services or data are hosted on remote infrastructure.

b. Ubiquitous: Services or data are available from anywhere.

c. Commodified: The result is a utility computing model similar to traditional that of traditional utilities, like gas and electricity - you pay for what you would want!

Fig.1 Cloud computing architecture

1.1 Allocation of Resource in Cloud Computing

In cloud computing, resource provisioning is an important issue as it dictates how resources may be allotted to an cloud application such that service level agreements (SLAs) of applications are met [2] .

Resource provisioning is based on FCFS scheduling algorithm, it analyzes response time distribution. This in turn is used to develop a heuristic algorithm, it defines an allocation scheme and it requires small number of servers. In responding to the effectiveness of the algorithm specification was evaluated in range of operating conditions[3]. A new modification called randomly dependent priority. It is originated to have the best performance in terms of required number of servers.

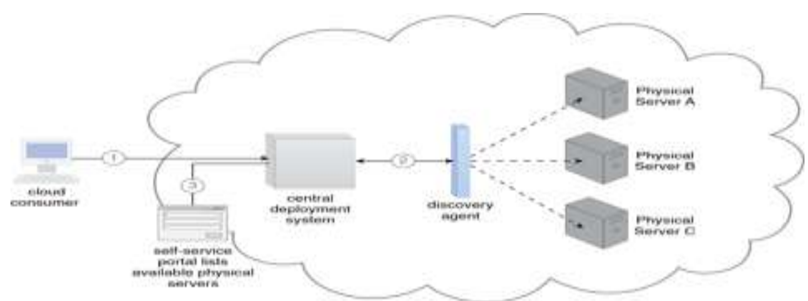

Fig.2 Allocation of resources in cloud environment

Five types of virtual instances are offered by Amazon EC2, each instance is attributed with different capacities in terms of RAM size, CPU capacity and I/O bandwidth [5] [9]. The declared volume details are virtual instances on EC2. To address fault- tolerance, EC2 distributes its virtual instances across several data centers, each data center is called as a availability zone [5][7]. To execute in different data centers, two virtual instances are running on different zones. There are six available zones, four are located in U.S. and the other two are in Europe. Further, to determine that the similar performance characteristics appear on different types of virtual instances as well, we also partially benchmark medium instances with high CPU.

1.2 Consequences of resource provisioning

As we noticed that not all small instances behave similar when serving CPU-intensive and disk I/O intensive workloads, we further notice this phenomenon and run the third group of experiment to verify if the CPU and disk I/O performances are interlinked on evidently matching small instances. Single virtual instance performances over CPU and IO are depicted by each point. Correlation between the respective CPU and I/O performances is not observed. To process different types of workloads, different types of small instances on Amazon EC2 may be apt [9] [10]. The above literature provided about discussion of the resource provisioning operations. We describe the efficiency in the resource provisioning of the cloud computing. An Enhanced optimal virtual machine Assignment (IOVMA) algorithm is proposed for providing the resources for VMs based on two provisioning schemes: reservation and on-demand. It can adjust the transaction between the advance reservation of resources and the allocation of on-demand resources. In addition, an IOVMA also takes the demand and price uncertainties into the resource provisioning. To further improve the IOVMA algorithm, authors of the IOVMA algorithm was proposed another optimal cloud resource provisioning algorithm, called the OCRP algorithm[6]. The OCRP algorithm extends the IOVMA algorithm to provision resources for VMs in multiple provisioning stages. To solve the optimal resource provisioning in an efficient way, two different approaches Benders decomposition and sample-average approximation are applied in the OCRP algorithm instead of the SIP model. For each VM, the placement information only indicates which cloud provider hosts the VM, not the information about the located PM.

1.3 Virtualisation

Virtual workspaces is an abstraction of an execution environment that can be made dynamically available to authorized clients by using well-defined protocols, Resource quota (e.g. CPU, memory share),Software configuration (e.g. O/S, provided services). It provide infrastructure API for Plug-ins to hardware/support structures

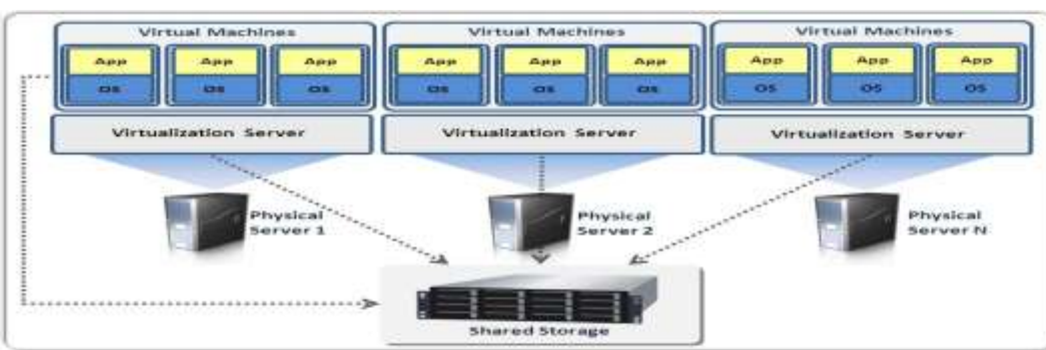

Fig.3 Virtualisation-Overview

Implementing Virtual Machines (VMs) as

- Abstraction of a physical host machine,
- Hypervisor intercepts and emulates instructions from VMs, and allows management of VMs,
- VMWare, Xen, etc.

II PRECEDING WORK

We define, the system mode that is used in this paper. Furthermore, we also review the previous studies related to our investigated Resource provisioning problem. For the VM placement issue, it has been discussed in a lot of literature. In the literature, the VM placement problem is usually transformed to the 0-1 knapsack (bin packing) problem. With the problem transformation, the PLP model corresponding to the VM placement can be easily formulated. Based on the derived PLP model, the optimal solution of the VM placement problem can be obtained [8] [10]. However, the previous VM placement literature focused on how to maximize the resource utilizations of PMs in the creation of VMs.

The amount of VM interference cost depends on various factors, such as the types of applications running in VMs, the number of VMs placed at the same PM, the choice of the VM scheduling algorithm.

2.1 PLP Representation

The PLP is a known mathematical method for solving the optimal problems with following characteristics: a linear objective function, a number of linear constraints, and an integer solution set. This model could be follow following assumptions. The cloud provider would like to create a number of new VMs in PMs concurrently [4] [5]. If the rent VM of a user cannot provide the computing environment to meet the QoS requirement of the user application, the cloud provide will return an amount of money to the user. Before placing the new VMs, each PM already has held a certain number of existing VMs. In the PLP model, after placing the new VMs in PMs, the primary function is to maximize the profit of the cloud provider.

According to the process of the virtual machine placement in commercial cloud computing, consider the cost approach for resource provisioning introduce IOVMA algorithm. Finally, VMs will be hosted in a computational environment which can be activated by third party sites termed as cloud providers. Regarding services of the cloud computing applications cloud service provider provides two types of planning services to the end users. Those services can be offered by the environment assurance in commercial cloud websites. Those are EC2, Go Grid, etc. These services are instance services and offer both reservation and on-demand plans to the end users. Generally, cost for resources in reservation scheme is less than that of on-demand scheme.

III BACKGROUND WORK

Mathematical translation of our problem statement is as follows: There are ‘M’ physical machines and the resource capacities of memory, CPU and Network bandwidth dimensions are provided. ‘N’ virtual machines are to be placed

[9] [11]. We need to identify mapping between VMs and PMs that satiates the VMs' resource requirements simultaneously curtailing the number of physical machines under use.

Resource demands are predicted at regular intervals using resource demand data. These predicted values are used by a placement module to compute VM to PM mappings. This module uses first fit approximation.Extracting each individual physical machines can be considered as bins having different dimensions in virtual machine placement. These dimensions are accessed in real time data processing with virtual machine object representation and other data elements along with client requirement specification in different dimensions. We have to define behavior of the each virtual machine placement with accurate resource generation. Resource allocation is the main achievement in optimized data delivery to clients according their requirements. Hence, due to the similarities of our problem with the bin -packing problem, we have adopted techniques like Superlative Linear Programming and First Fit which are typically used to solve traditional bin-packing problems, to solve our problem of VM Placement.

IV ALGORITHM ANALYSIS

Inputs for the algorithms include resource needed by each virtual machine to be allocated. We define a requirements matrix in order to capture these requirements through various placements with virtual organizations. The matrix is described as follows:

Requirements Matrix: {r11, r12 …r1d}.

Consider the above representation where each Rij represents the requirement of VM i long as the dimension j.

Currently, we consider three dimensions for our purposes: CPU, Memory and Network bandwidth use [4] [5] [6]. Requirements along these dimensions are expressed as fractions of the total capacity of a PM.

4.1 Superlative Linear Programming

Although we believe that this particular prototype is valuable, there are three important knapsacks: the storage options under each dataset is not systematic, that is, we are not allowed to specify a preferred storage option; since each dataset is analyzed in isolation the best global solution is not obvious; and the computational specifications (number of application runs, cost per hour, machine speed, etc.) are not considered.

Insert Equation $(x+a)^n = \sum_{k=0}^{n} \binom{n}{k} x^k a^{n-k}$

Basically, the cloud provider would like to make PMs hold VMs as many as possible to generate more revenue. As increasing the number of VMs in a PM, it may introduce more interference among the VMs in the same PM. This will possibly increase the penalty payment of the cloud provider in the VM provisioning.

We are making an attempt to address these knapsacks so that we can obtain a solution to global data allocation that balances the cost and performance of both storage and computation [11]. If the best storage service for a single dataset that reside in the cloud does not have good options for the rest of the application data, then we may arrive to a sub-optimal allocation of data. As we will show, trying to find an optimal solution, that convolutes the problem to NP hard. We try to provide a ubiquitous model for this data allocation problem and a software implementation that is both fast and scalable.

V PERFORMANCE ASSESSMENT

The Enhanced Optimized Virtual Machine Assignment performed over the following parameter settings. In each PM, there have been a number of existing VMs in it. The number of the existing VM is randomly determined from 0 to 10 [5] [7]. With holding different existing VMs, the amount of available resources in each PM is also different. The amount of available resources is represented as a triple-tuple (available CPU GHz, available memory space in GB, available storage space in GB). The resource interval [(12, 129, 200), (96, 3000, 9600)] is used to randomly decide the available resources of each PM. In each PM, it uses a 40Gbps transmission line to connect with the corresponding switch. Next, a number of new VMs is assumed to be created within 250 PMs. The number of new VMs is set from 100 to 500 in each simulation run, respectively. The amount of the resources required for a new VM was set by referring to the Amazon EC2 with 12 different resource demands [7] [8] [9]. In simulation experiments, we also refer to Amazon EC2 to set the price of each VM type and the QoS requirement of an application running in a VM. If the QoS violation is decided, the penalty payment is set using the violation ratio × the price of the VM.

For the number of VMs created in the PMs, all the four algorithms have similar simulation results. Basically, the least fit algorithm can fully exploit the resources of PMs since it attempts to place the VM at the PM with few resources. Therefore, the least-fit algorithm should have better performance in the number of VMs created than the other two intuitive algorithms and our proposed heuristic algorithm. In the proposed OVMP algorithm, it also attempts to place many VMs to reducing the VM interference for maximizing the profit of the cloud provider.

VI EXPERIMENTAL RESULTS

We define the resource provisioning in cloud with description of all the resources. Compute or design all the relations of cloud computing.

6.1 Evaluation of Resource Allocation

In this evaluation process of extraction of various applications in cloud resource provisioning operations.

6.2 Cost Estimation

Based on the various operations present in the cloud computing environment, We describe the cloud service provider by analyzing the IOVMA algorithm with suitable consideration on reservation and on-demand cost estimation in real time virtual machine placement and in real time cloud environment. Yet, reserving too many VMs may not be optimal. Therefore, the balance between on-demand and oversubscribed costs needs to be obtained in which IOVMA performance is optimal.

6.3 Implementation

To represent long term planning we are using the IOVMA algorithm that can be applied to multiple provisioning stages. Optimal solution of the first provisioning stage depends on multiple probability and randomly distribution with consideration of occurring in sequential time operations. Multiple stages with planned and achievement releases with suitable examples. For example consider the systematic data events dynamically high efficiency various time periods in a year (e.g., Christmas Day, Valentine's Day, etc.) [2][9]. The use of decomposition method for IOVMA has to be carefully considered, since the formulation of the IOVMA algorithm is a pure integer program which is the NP-hard problem. Even though the sub problems can be solved simultaneously, the master problem with the additional Benders cuts requires considerable computational time.

An Enhanced optimal virtual machine Assignment(IOVMA) algorithm is proposed in order to reduce the total cost due to buying reservation and on-demand plans of resource provisioning. IOVMA algorithm helps to make a decision to host a certain number of VMs on appropriate cloud providers for providing IAAS. Randomness of future demands and prices of resources are considered to optimally adjust the tradeoff between on-demand and oversubscribed costs [10] [11]. The decision made by IOVMA algorithm is obtained as the optimal solution from Superlative integer programming (EIP) formulation with two- stage recourse. Extensive derivation of mathematical models and simulation in cloud computing environment are performed to examine the effectiveness of IOVMA algorithm. The results show that IOVMA algorithm can minimize the total cost, while satisfying requirements of both providers and customers.

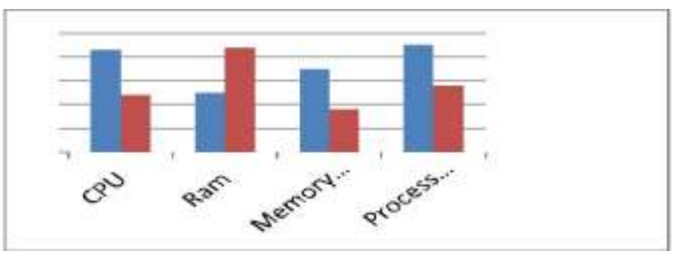

Fig. 4 Simulated result of comparison with OCRP & IOVMA

The OCRP formulates the significance of the other processors and other data management services take same environmental situation [7] [8]. the loading time other institutions are achieved in real time data passing between operating services present in cloud computing. The IOVMA procedure gives efficient and excellent improvement of the resource provisioning processing of resource like CPU and other devices present in cloud computing operations

VII CONCLUSION

Solving Superlative integer programming with multistage recourse will result in the optimal solution. PIP is obtained from OCRP. Benders decomposition approach to divide an OCRP problem into sub problems which can be solved parallels is also applied. An Enhanced optimal virtual machine placement (IOVMA) algorithm is proposed in this paper, this algorithm can be extended to implement optimized resource provisioning operations. This algorithm could reduce the cost spent in each plan for hosting virtual machines in a multiple cloud provider environment with future demand and price uncertainty.

REFERENCES

[1] Introduction to Cloud Computing, Fact Sheet, Fiche d'information

[2] "Resource Provisioning for Cloud Computing", IEEE International Conference on Autonomic Systems, Barcelona, 2009, Ye Hu1, Johnny Wong, Gabriel Iszlai, and Marin Litoiu,

[3] "Efficient Resource Provisioning in Compute Clouds via VM Multiplexing", ICAC'10, June 7–11, 2010, Washington, DC, USA. Copyright 2010 *ACM 978-1-4503-0074-2/10/06 ...$10.00*, Xiaoqiao Meng, Canturk Isci, Jeffrey Kephart, Li Zhang, Eric Bouillet,

[4] "SLA-Oriented Resource Provisioning for Cloud Computing: Challenges, Architecture, and Solutions ", Rajkumar Buyya, Saurabh Kumar Garg, and Rodrigo N. Calheiros, 2011 International Conference on Cloud and Service Computing, 978-1-4577-1637-9/11/$26.00 ©2011 IEEE.

[5] "Optimization of Resource Provisioning Cost in Cloud Computing", Sivadon Chaisiri, Bu-Sung Lee, IEEE Transactions On Services Computing, Vol. 5, No. 2, April-June 2012.

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

[7] "Profiling Energy Usage for Efficient Consumption," R. Chheda, Stokowski, S. Stefanovich and J. Toscano, Architecture J., no. 18, 2008. "SLA-Aware Virtual Resource Management for Cloud Infrastructures," H.N. Van, F.D. Tran, and J.-M. Menaud, Proc. IEEE Ninth Int'l Conf. Computer and Information Technology, 2009.

[8] "The FEDERICA Project Website," http://www.fp7-federica.eu, 2013.

[9] D. Andersen, "Theoretical Approaches To Node Assignment,"Unpublished Manuscript, http://www- . cs.cmu.edu/~dga/papers/andersen-assign.ps, 2002.

[10] M. Yu, Y. Yi, J. Rexford, and M. Chiang, "Rethinking Virtual Network Embedding: Substrate Support for Path Splitting and Migration," *ACM SIGCOMM Computer Comm. Rev., vol.38, no. 2,pp. 17-29, Apr. 2008, doi:10.1145/1355734.1355737.*

[11] W. Szeto, Y. Iraqi, and R. Boutaba, "A Multi-Commodity Flow Based Approach to Virtual Network Resource Allocation," Proc.IEEE GLOBECOM '03, vol. 6, pp. 3004-3008,Dec.2003, doi:10.1109/GLOCOM.2003.1258787.

[12] J. Fan and M.H. Ammar, "Dynamic Topology Configuration in Service Overlay Networks: A Study of Reconfiguration Policies," Proc. IEEE INFOCOM '06, pp. 1-12, Apr. 2006, doi:0.1109/INFOCOM.2006.139.

[13] J. Lu and J. Turner, "Efficient Mapping of Virtual Networks onto a Shared Substrate," Technical Report WUCSE-2006-35, Washington Univ. St. Louis, 2006.

[14] Y. Zhu and M.H. Ammar, "Algorithms for Assigning Substrate Network Resources toVirtual Network Components," Proc. IEEE INFOCOM '06, pp. 1-12, Apr. 2006, doi:10.1109/INFOCOM. 2006.322.