Bio-inspired computation for task scheduling in hybrid cloud

M.Anuradha, Dr.S.Selvakumar

Abstract—IaaS(Infrastructure as a Service) is a type of cloud computing in which a third party provider hosts virtualized computing resources over the internet. IaaS achieves economy of scale by providing services to multiple clients at one time called multiplexing. Number of scheduling approaches have proposed in the recent years but requires standardization and agreement hence creates difficulties for the cloud users. The proposed system addresses the issue by sending task to other providers and utilizes the unused resources by analyzing it through Ant Colony Optimization(ACO). The intelligent scheduler in the proposed approach schedules task in an optimized manner in the way how ants search food from its ant hill.

Index Terms— Ant Colony Optimization(ACO), IaaS, resource, scheduling, task,provider.

1 Introduction

Cloud computing is a subscription-based service where a network storage space and computation resources can be obtained. IaaS(Infrastructure as a Service) has become an important paradigm in cloud computing, it achieves economy of scale by multiplexing and therefore faces the challenge of scheduling task. The proposed scheduling approach, based on ACO (Ant Colony Optimization) schedules task from private cloud to external cloud when there is a peak demand. The system is modeled as a framework where a provider outsources its tasks to external cloud (Ecs) to meet the unpredictable situation in cloud environment. The approaches so far used either rely on inter-cloud agreement that is required for the cloud federation. The proposed scheduler allocates the task to resource and adopts dynamic change in the cloud environment without any standardization also it provides optimized scheduling by which the factors of a good scheduler such as high throughput and reduced response time can be achieved and solves the issue of portability while migrating to different cloud.

Infrastructure as a Service (IaaS) becomes very popular as the foundation for higher level services such as Platform as a Service (PaaS) and Software as a Service (SaaS) [1]. IaaS providers such as Amazon EC2 and IBM Smart Cloud Enterprise [2], allow users to rent resources in the form of Virtual Machines (VMs). They can offer different VM types that are characterized by machine configuration, QoS[10] and pricing model.

A straightforward solution for a cloud provider is to over purchase the cloud resources or relying on standardization and agreement as in cloud federation.

- *M.Anuradha is currently pursuing masters degree program inComputer science and engineering in GKM College of Engg and Technology., Chennai, India.PH-9551674174. E-mail: anuradha.sairishi@gmail.com.*
- *Dr.S.Selvakumar received Doctor of philosophy in computer science and engineering from anna university,Currently working as a Professor and Head in GKM College of Engg and Technology. Chennai,India. PH-9789916648. E-mail:sselvakumar@ieee.org*

The proposed system provides a framework in which the IaaS provider can utilize the unused resources in one or more public cloud called External cloud(ECs) to overcome the resource limitation.

IaaS provider acts like a private cloud where as the task are outsourced to public cloud called External clouds(ECs).

2 Motivation for Proposed Approach

There exists no generic model for task scheduling in cloud computing environment[16] because most of the time the user requirements are unpredictable. Especially, mapping of user specified configuration to the available resource level with the provider is difficult also modeling non-functional requirements such as availability and performance are challenging.

Parameterization provides an efficient way for analyzing the solution but modeling parameterization is difficult when the problem size is large and dynamic which is usually a case in cloud task scheduling where there exists user traffic for submitting task and getting back their results.

The cloud environment for task scheduling requires to be continuously optimized. Given the dynamic nature of clouds, resource allocation needs to be updated continously for the QoS requirement . It is challenging to efficiently decide when and how to reconfigure the cloud in order to dynamically adapt to the changes.

3 Modeling Of Components

The cloud environment is modeled as different layers in the Consumer layer the users of a cloud are modeled to create user traffic. The requests submitted as task by the user is sent to the business layer through the presentation layer where actual scheduling is carried out which constitutes private cloud i.e., a cloud service provider.

The user of a cloud not only interacts with user interface and also the business logic by submitting their requirements through configuration.The submitted requests are processed by the request handler and is sent to the scheduler. The scheduler gets the availability of resources through the re-

source monitor which interacts with the resource pool containing datacenter of cloud encapsulates hosts which in turn provides virtual maching based on user requirement.

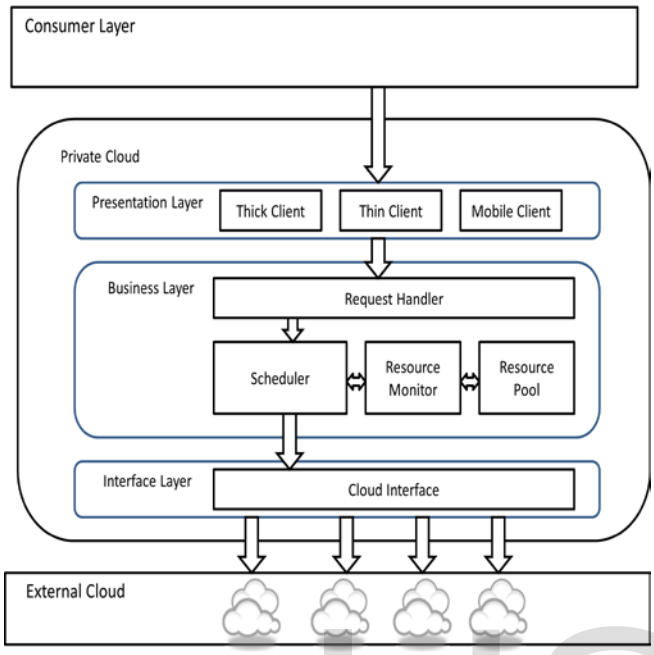

Fig 1. Component view of proposed framework

4 Solution Framework

Consider[18] CP={CP_1,CP_2,…,CP_n} be the set of cloud providers. Assume CP_1 is the private cloud and CP_2,…,CP_n are external clouds. VM={VM_1,VM_2,…,VM_I} be the set of VM types and A={a_1,a_2,…,a_w} be the set of applications required to be scheduled for each VM and each application has a runtime rj and a task set {t_{j1},t_{j2},…,t_{jIj}}.

An Integer programming model is formulated to solve this problem. The objective of integer programming formulation is to maximize the profit of private cloud on the premise of guaranteeing QOS. To formulate this problem, problem parameters and decision variables are defined in Table 1 and Table 2, respectively.

The problem can be formulated as the following IP model. Maximum Profit achieved by the cloud service provider is given by

$$\text{Profit} = \sum_{j=1}^{w}\sum_{v=1}^{I} T_j b_{jv} p_v r_j - \sum_{j=1}^{w}\sum_{l=1}^{T_j}\sum_{v=1}^{I}\sum_{k=1}^{n} y_{jlk} b_{jv} c_{kv} r_j$$

The first term of objective function represents the income of CP1 and the second one means its cost[17]. Constraint specified in the formulation guarantees that each task is allocated to exactly one cloud provider. The problem is formulated as an IP model for task allocation. Solving such problems using a mathematical programming approach will take a large amount of computational time[18].

TABLE 1

PROBLEM PARAMETERS

n	Number of cloud providers
I	Number of VM types
w	Number of applications
p_v	Price of the vth VM type in CP_1
c_{kv}	Cost of the vth VM type in CP_k
d_j	Deadline of the jth application
r_j	Runtime of each task in the jth application
T_j	Number of tasks in the jth application
b_{jv}	If $b_{jv}=1$, the jth application use VM type VM_v; otherwise, it does not use this type.

TABLE 2

Binary decision variable

y_{jlk}	Binary decision variable, such that y_{jlk}=1 if t_{jl} is allocated to CP_k; otherwise $y_{jlk}=0$

From the formulation, it is observed that solving such problems using mathematical programming approach will take large amount of computational time for a huge task scheduling problem in cloud environment. Hence there is a requirement for Integer programming model which promotes for a real time scheduling.

5 Ant Colony Optimization

Ant colony optimization is a probabilistic technique useful in problems which deal with finding better paths through graphs. Artificial or simulated agents locate optimal solutions by moving through a parametric space representing all possible solutions.

Natural agents lay down pheromones directing each other to food source while exploring the environment. Similarly, ants record their positions and the quality of their solutions while migrating for resources.

So that in later iterations more ants locate better solution by knowing the history of their predecessors. Here, simulated agents are considered as user's task and the food source is cloud resources such CPU, Memory, bandwidth etc., The role of ACO in the proposed system is to schedule task in an optimized manner.

Real ants are deaf, dump and blind it communicates to its companion about the food source by laying a substance called pheromone on ground. When more ants follows the path the

ejected substance form a trail through which other ants are attracted and hence communicated.

The more the ants more denser the pheromone hence more followers finds food in a shortest path.In the similar way the proposed algorithm works for finding suitable resources for the user submitted task and finds an optimal solution. The pheromone value determines how fast the virtual machine performs comptation in the cloud environment.

6 Steps for ACO Datacenter broker

Initialization
- Setting VM time usage to zero
- Idle VM to available

Updating pheromone table
- Based on the finishing time of the task in a VM
- Updates pheromone value on each trail

Finding the optimal VM
- Identify the VM with minimum task execution time.
- Assign the task.

Finishing time
- Setting the minimum time used by the VM for task execution.

Ant tour
- Select the task in radom
- Assign the task to VM with maximum capability
- Finds the optimal solution based on finish time
- Updates pheromone value

Finding solution
- Perform ant tour
- Output the record
- Identify the solution based on finish time
- Set the solution as optimal solution
- Output the pheromone value

7 Implementation Using CloudSim

CloudSim is a framework for modeling and simulating cloud computing infrastructures and services developed in CLOUDS laboratory at the university of Melbourne.

7.1 Why CloudSim?

1) Models and simulates large scale cloud computing data centers.
2) Models and simulates virtualized server hosts with policies for provisioning host resources to VMs.
3) Inserts simulation elements dynamically, stops and resumes simulation.
4) Allows user-defined policies for the allocation of host to VMs and for allotting host resources to VMs.

By making use of the provision for extending the user-defined policies the ACO Datacenter broker policy is extended from Datacenter broker of CloudSim.

7.2 Implementation Results

The Execution time required for scheduling using ACO Datacenter policy is lesser when compared with the existing policy. The use of ACO utilizes the resources better than the existing Roundrobin policy implemented in CloudSim.

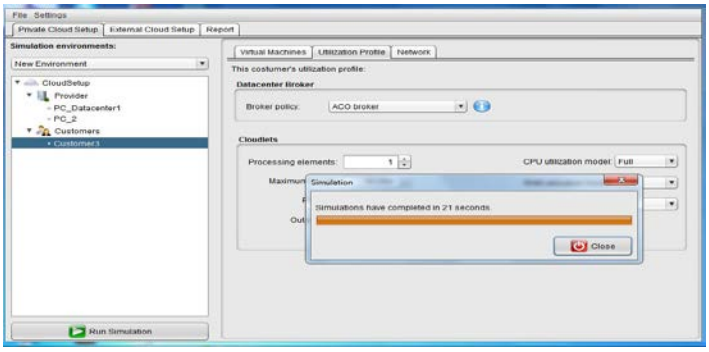

Fig 2. Execution time for ACO based broker

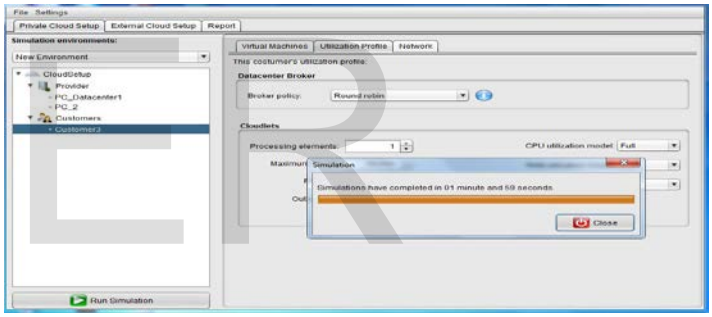

Fig 3. Execution time for Round robin

8 Performance Evaluation

Experiments were conducted using CloudSim[3], the cloud simulator. Initially user traffic were generated for the allocation of requests in private cloud. Round robin based scheduling approach were used for the scheduling in the private cloud.

In order to show insufficiency of resources in private cloud more number of unpredictable requests were generated .ACO based scheduling is performed among the public cloud for the suitable and optimal allocation of user requests.

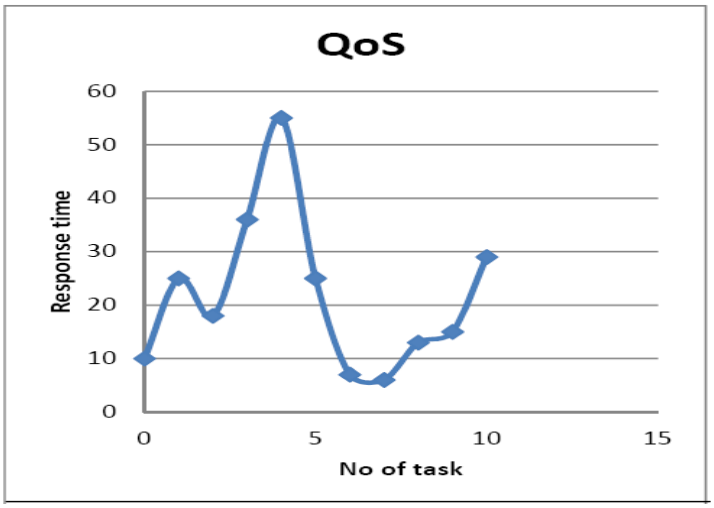

Fig 4. Response of a provider for a number of tasks

As a result of simulation we observed that the proposed system improves the profit of cloud provider by satisfying the user who could not understand where their actual tasks are executed. And at the same time it provides QoS to the user where no user requests are rejected irrespective of the situation.

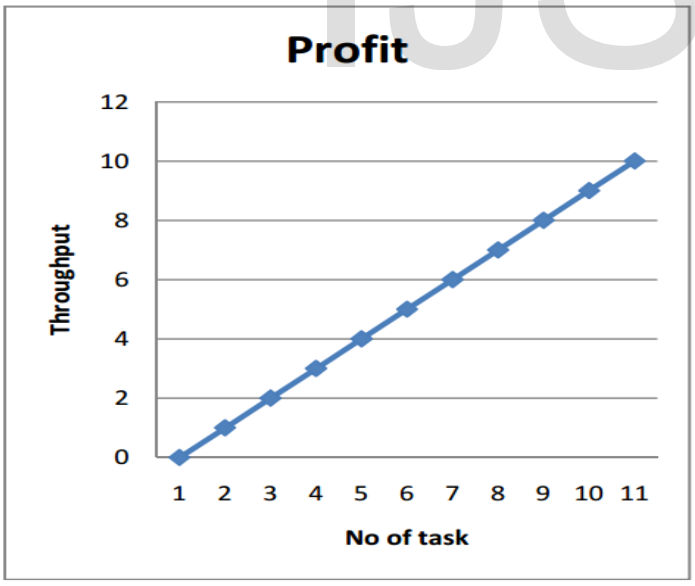

Fig 5. Profit of the provider is increasing

9 Conclusion

The proposed system addresses the issue of scheduling task in cloud environment by providing a framework in which a cloud provider outsourses its request to External clouds in order to utilize the unused resources and to increase the productivity.Thus the proposed approach based on ACO performs intelligent scheduling and achieves more profit for the cloud service provider and provides user with the better utilization of resources based on their need.

References

[1] Cui Lin, Shiyong Lu, "Scheduling Scientific Workflows Elastically for Cloud Computing" in IEEE 4th International Conference on Cloud Computing, 2011.

[2] Dar-Tzen Peng, Kang G. Shin and Tarek F. Abdelzaher, " Assignment and Scheduling Communicating Periodic Tasks in Distributed Real-Time Systems", IEEE Transactions On Software Engineering, Vol. 23, No. 12, pp. 745-758, 1997.

[3] Dr. Rajkumar Buyya, "CloudSim: a toolkit for modelling and simulation of cloud computing environments and evaluation of resource provisioning algorithm", published online 24 august in Wiley Online Library 2010, pp. 23-50.

[4] J. M. Hyman, J. M. Hyman, A. A. Lazar, A. A. Lazar, G. Pacifici, and G. Pacifici."Real-time scheduling with quality of service constraints" IEEE Journal on Selected Areas in Communications,, pp:1052–1063, 1991.

[5] K. Thirupathi Rao, P. Sai Kiran, L.S.S Reddy, V. Krishna Reddy, B. Thirumala Rao, "Genetic Algorithm For Energy Efficient Placement Of Virtual Machines In Cloud Environment", in proc IEEE International Conference on Future Information Technology (IEEE ICFIT 2010), China, pp:V2-213 - V2-217.

[6] Kennedy, J.; Eberhart, R., "Particle Swarm Optimization", IEEE International Conference on Neural Networks, Perth, WA, Australia, Nov.1995, pp.1942-1948.

[7] L. M. Gambardella and M. Dorigo. "An ant colony system hybridized with a new local search for the sequential ordering problem" INFORMS Journal on Computing, 12(3):237–255, 2000.

[8] M. Dorigo and T. St¨utzle. "Ant Colony Optimization" MIT Press, Cambridge, MA,2004.

[9] M. Dorigo, V. Maniezzo, and A. Colorni, "The ant system: an autocatalytic optimizing process ", *Technical Report TR91-016*, Politecnico di Milano (1991).

[10] Meng Xu, Lizhen Cui, Haiyang Wang, Yanbing Bi, "A Multiple QoS Constrained Scheduling Strategy of Multiple Workflows for Cloud Computing", in IEEE International Symposium on Parallel and Distributed Processing, Chengdu 10-12 Aug,2009, pp 629-634.

[11] Strategy of Multiple Workflows for Cloud Computing", in IEEE International Symposium on Parallel and Distributed Processing, Chengdu 10-12 Aug, 2009, pp 629-634.

[12] Strategy of Multiple Workflows for Cloud Computing", in IEEE International Symposium on Parallel and Distributed Processing, Chengdu 10-12 Aug, 2009, pp 629-634.

[13] T. D. Braun, H. J. Siegel, N. Beck, L. L. B¨ol¨oni, M. Maheswaran, A. I. Reuther, J. P.Robertson, M. D. Theys, B. Yao, D.Hensgen,andR.F.Freund."A comparison of eleven static heuristics for mapping a class of independent tasks onto heterogeneous distributed computing systems" *Journal* of Parallel and Distributed Computing, 61:810–837, June 2001.

[14] Tzu-Chiang Chiang, Po-Yin Chang, and Yueh-Min Huang,"Multi-Processor Tasks with Resource and Timing Constraints Using Particle Swarm Optimization", IJCSNS International Journal of Computer Science and Network Security, Vol.6 No.4, pp. 71-77, 2006.

[15] V. Krishna Reddy, B. Thirumal Rao, L.S.S. Reddy, P. Sai Kiran "Research Issues in Cloud Computing" Global Journal of Computer Science and Technology, Volume 11, Issue 11, July 2011

[16] Ruben Van Den Bossche, Kurt Vanmechelen and Jan Broeckhove,"Cost-Optimal Scheduling in Hybrid IaaS Clouds for Deadline Constraied Workloads",IEEE 3rd International Conference on Cloud Computing.2010.

[17] Xingquan Zuo, *Member, IEEE*, Guoxiang Zhang, and Wei Tan, "Self-Adaptive Learning PSO-Based Deadline Constrained Task Scheduling for Hybrid IaaS Cloud.", IEEE Transactions On Automation science and engineering, Vol.11,No.2,April 2014.

[18] Zhangjun Wu, Xiao Liu, Zhiwei Ni, Dong Yuan, Yun Yang, "A Market-Oriented Hierarchical Scheduling Strategy in Cloud Workflow Systems" in JSC2010.