International Journal of Emerging Technology in Computer Science & Electronics (IJETCSE)
ISSN: 0976-1353 Volume 21 Issue 1 – APRIL 2016.

MULTI-OBJECTIVE OPTIMAL RESOURCE ALLOCATION USING IMPROVED ACO

K.Soundarya[#1] and R. Jayamala[*2]

[#] *M.E, Mobile and pervasive Computing, University College of Engineering-BIT campus, Trichy, India*

[*] *Asst.professor , Information Technology, University College of Engineering-BIT campus, Trichy, India*

***Abstract*— The cloud systems are highly unpredictable with resource availability. Sometimes single cloud can't able to handle widespread data loss, downtime due to failure, the huge amount of data and sudden increase in demand. Therefore multi-cloud becoming popular nowadays. Optimally allocating resource based on the demand and availability is always an NP-complete problem. The more optimized structure of algorithm is needed for resource sharing, otherwise, the resource utilization is reduced and also the response time will get increased. So, in this paper, an Improved ACO (Ant Colony Optimization) is used for resource allocation. The resources are allocated to task with minimum cost, time, power and energy. Therefore, both the user and provider will get benefit out of it.**

Index Terms— Multi-cloud, Ant Colony Optimization, resource allocation.

I. INTRODUCTION

Before cloud computing, the applications were executed on a specific system. We have to install a suitable software for each computer instead of that only have to load one application. That application allows anyone to access the service anywhere through the internet. The Remote machine owned by some company process complex data and analyze it. It will charge the users based their demand (Pay-as-you-go). This is called Cloud Computing. Through any terminal equipment, the stored data can be accessed anytime by the user using API (Application Programming Interface) provided by cloud provider with the help of internet. Even though, the cloud having numerous advantages like eliminating the server and software investment, backup and recovery still there are several challenges in the single cloud like a vendor-lock-in problem (migrating service from one cloud to another), Downtime (component failures).To resolve these challenges we moving towards Multi-Cloud. The multi-cloud not only to offer hardware, software or to optimize fault tolerance. It also steers the traffic from various users to the fastest possible path of the network. For a particular task, some clouds are better suited than others. For example, some provider might handle large numbers of requests per unit time using small data transfers on the average, but another provider might perform better for smaller numbers of requests per unit time using large data transfers on the average.

Multi-cloud major advantages:

- Scalability-- To address peak demands
- Collaboration--Sharing of infrastructure between partners
- Deployments In Multi-Site- Aggregation of Infrastructure across distributed data center
- Reliability--Fault tolerance architectures
- Performance- Service Deployment closer to end users
- Cost--Dynamic placement reduce the overall infrastructure cost
- Energy Consumption--Min energy consumption

Scheduling workloads in the multi-cloud environment are challenging as the resources in the data centers are heterogeneous in nature. The requirement of each task may vary. Some task requires more CPU resource other may request more storage and the cost also different for different resources. Efficient scheduling must adapt various environment. Even though the resources are vast and reliable in the cloud, dynamically managing and allocating resources across different cloud is a crucial problem. To overcome these we need an optimization algorithm that schedules the task effectively to the resource available. In general, ACO (Ant Colony Optimization) will converge easily to a local optimal solution. This paper introduces a multi-objective improved ACO to solve the above issue. The multi-objective problem will not consider one best solution, instead of that it will consider a set of solutions taking all objective. This set is called as a non-dominated solution or Pareto set.

The rest of the paper organized as follow: some work related to the topic is given in section II. The system model and problem definition are discussed in section III and IV respectively. The Improved Multi-objective based ACO algorithm in section V and finally conclusion in section VI.

II. RELATED WORK

The work in [1] solves big work flow scheduling and multiple QOS parameters using ant colony optimization. This allows users to specify their own preferences for service and QOS. It also consider budget price constraints. However, this method is used for scheduling workflow in grid computing. Based on the shortest delay time of tasks, [2] proposes an improved ant colony algorithm consideration both equity and efficiency. The paper [3] proposes a scheduling strategy based on ant colony optimization and a two-way mechanism. It sets a pheromone threshold to avoid premature convergence and using the pre-execution time it also avoids local optimum through the two-tier search strategy. Xia-yu Hua, Jun Zheng, Wen-Xin Hu [4] introduce a resource allocation method for cloud based on ACO. To predict the execution speed it fully consider the node's load and the inherent properties of computing resources. Dorigo, M.; Maniezzo, V.; Colorni, A, In [5] discussed Ant Colony Optimization (ACO). It's a general-purpose heuristic algorithm, which can be used to solve the different combinatorial optimization problem. Also, it's a population –based approach in which parallelization can easily be achieved.In [6], the ACObased approach is introduced as an instance of the multidimensional bin-packing problem. However, this method is modeled on a single-objective method for minimizing the number of physical machines. In [7] the MOTS (Multi-objective Task Scheduling) algorithm is used to resolve the tradeoff between makespan (Total execution time) and cost but optimization problem is not considered. A Dynamic Optimization Algorithm for Task Scheduling in Cloud Environment: Monika Choudhary and Sateesh Kumar Peddoju presented this paper [8]. They merged the concepts of three different approaches i-e Task Grouping, Prioritization and Greedy Allocation which permits improved performance and provide new way of scheduling. It works on two constraint groups deadline constrained and cost based tasks. Task broker groups the tasks according to above constraint categories after it received all the task, and it assign priority to each task based on their group type.

III. SYSTEM MODEL

Multi-cloud is nothing but a set of interconnected cloud. The VM's are of different capacity and type. The data centers may not be sufficient. So, collaboration of multiple data centers may require to solve scientific and engineering applications. We assuming that manager server in each cloud will track all the detail about that particular cloud and communicate this information with other cloud manager in order through a global manager to divide the workload effectively. The user submit the task (*Ci*; *Mi*; *Di*; *Li*) to the task manager, it collect all the detail and submit the request to the global manager through scheduler. The global manager collecting and updating the information about each cloud periodically. It also get the input (Provider offerings, resource status and available resource) from the service provider. The Global manager will compute resource requirement and power consumption of each cloud and schedule the task based on the multi-objective ACO solution.

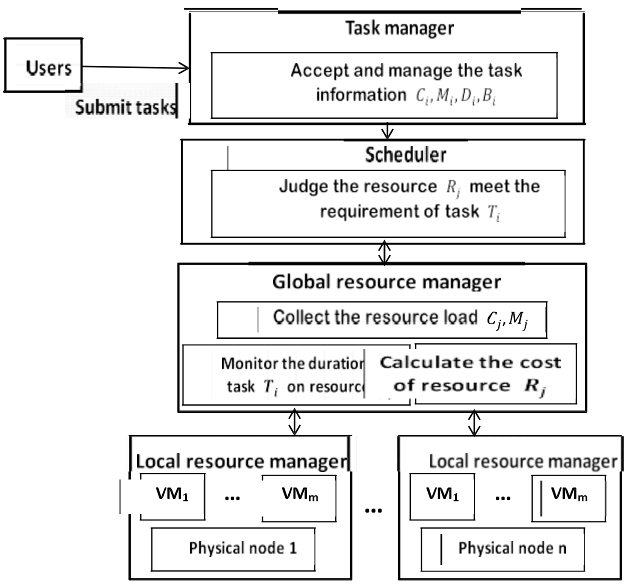

Figure 1: system model

TABLE I. INPUT PARAMETERS

Parameter	Provider	Customer
Provider Offerings	√	-
Resource Status	√	-
Available Resources	√	-
Application Requirements	-	√

IV. PROBLEM DEFINITION

The integration of cloud provider activities for allocating scarce resources and utilizing it within the limit of cloud atmosphere so as to satisfy the needs of cloud application is called Resource Allocation Strategy (RAS).An optimal RAS should avoid resource contention, scarcity of resource, over and under provisioning, and resource fragmentation. The Problem is to map set of K $\{K_1,\ldots,K_k\}$ task from user to set of R $\{R_1,..,R_n\}$resource in set of $C\{C_1,..,C_c\}$ cloud based on the multi-objective optimization method, so that it will be benefit to both the provider and user by reducing cost, makespan, power and energy consumption.

minimize $\sum_x H(x) = F(x), B(x)$ (a)
s.t $B(x) = C(x) + M(x)$ (b)
$B(x) \le \sum_{i=1}^{k} B_i$ (c)
$F(x) \le \sum_{i=1}^{k} D_i$ (d)

Here *x* is a feasible solution. *F*(*x*) is a function of the performance objectives that refer to makespan. *B*(*x*) is the objective function of the user budget costs which consist of the task demand costs for the CPU and memory that correspond to the cost functions in the resource cost model. This is a multi-objective optimization problem that is difficult to solve; in particular, it is difficult to obtain the optimal solution. Consequently, this paper uses the ant colony algorithm to solve this problem.

Notation with Definition:

C_i, M_i - CPU and Memory of Task T_i	T_i - Task i
R_j - Resource j	$C_{cost}(j), M_{cost}(j)$- CPU and Memory of resource R_j
C_{base} - The base cost of CPU under the lowest usage	M_{base} - The base cost of memory
α , β - The weight factor of heuristic information and pheromone	t_{ij} - The duration time of T_i in resource R_j
C_{Trans}, M_{Trans}- The transmission cost associated with CPU and memory	ℓ - Pheromone evaporation factor
ɣ ,δ – Weight factor of performance and cost	U^p_j - CPU Utilization

P_j^{busy}, P^{ideal}_j -the average power values when the j-th PM is busy and idle, respectively	
PM – Physical Machine	
VM- Virtual Machine	

V. IMPROVED MULTI-OBJECTIVE BASED ACO

Here, the task $Ti = (Ci;Mi;Di;Li)$ where $Ci;Mi;Di;Li$ are the cpu usage, memory usage, deadline of the task and budget cost of the user respectively. These parameter submitted by the user to the task manager. The resource is defined by $Rj = (Cj;Mj)$.these two parameter are cpu utilization and memory usage.before proceeding further the assumption we made is user will submit only trusted information i.e. the request for resource is accurate.

The resource cost is divided into two parts cpu and memory.

$C_{cost}(j) = C_{base} \times C_j \times t_{ij} + C_{Trans}$ (1)

$M_{cost}(j) = M_{base} \times M_j \times t_{ij} + M_{Trans}$ (2)

Therefore total cpu and memory cost is given by

$C(j)= \sum_{j=1}^{N} C_{cost}(j)$ (3)

$M(j)= \sum_{j=1}^{N} M_{cost}(j)$ (4)

When an ant traverses all tasks, forms a path. This path must be feasible solution to the problem. The main two target is makespan and cost reduction considering that as fitness function $F(x)=\gamma e^{(-F(x))} + \delta e^{-B(x)}$ (5)

The ant choose the path randomly. When it reach their target it will calculate the path fitness. Later based on the probability calculated the ant will proceed further and it will focus on the high fitness path based on the phrenome value.

$$P_k(T_i, R_j) = \begin{cases} \frac{[\tau(T_i,R_j)]]^{\alpha}[\eta T_i,R_j)]^{\beta}}{\sum_{h \varepsilon g_k (T_i,R_j)} [\tau(T_i,h)]^{\alpha}[\eta[(T_i,h)]^{\beta}} & R_j \varepsilon\ g_k(T_i, R_j) \\ 0 & otherwise \end{cases} \quad (6)$$

$$\Delta\tau(T_i, R_j)=\begin{cases} Q\gamma e^{(-F(x))} + \delta e^{-B(x)} & ,(T_i, R_j)\varepsilon \text{ Path} \\ 0 & ,otherwise \end{cases} \quad (6a)$$

The solution will optimized locally, to avoid that we use pheromone evaporation and it's necessary to update the behaviour each time.

it is necessary to update the pheromone for each side of the path.

$\tau(T_i, R_j) =(1-\ell).\ \tau(T_i, R_j) +\Delta\ \tau(T_i, R_j)$ (7)

$\Delta\ \tau\ (Ti; Rj)$ is the incremental amount of the pheromone.

The global resource manager compute (8) to minimize the energy consumption of a data center

$P_j=(P_j^{busy}-P^{ideal}_j) * U^p_j + P^{ideal}_j$ (8)

To minimize resource wastage. Assumed that W is the resource wastage of each PM and R^{PM} indicates the set of resources available in PM^j Also $R_{1,I}^{VM}$,presents the set of CPU resources requested by VM_i

$$\min \sum_{j=1}^{M} W_j^{PM}=\sum_{j=1}^{M}[y_j.(\frac{R_{j,1}^{PM}-\sum_{i=1}^{N}(x_{ij}.R_{i,1}^{VM}}{\sum_{i=1}^{N}(x_{ij}.R_{i,1}^{VM}})] \quad (9)$$

ALGORITHM STEPS:

Step1: Get relevant parameter as input like number of task, resource required, task deadline, budget, and availability.

Step 2: The each cloud ant randomly choose the path. Each ant assigned to task. The task Ti when successfully allocated to resource Rj it is recorded in the table.

Step 3: In each next iteration until obtaining optimal solution the ant choose task based on the probability (6)

Step 4: The steps are repeated until all the task are scheduled completely, in each cloud

Step 5: Update the pheromone value locally which is best fit and update that globally.(6a)

Step 6: The good solution will get enhanced and other get eliminated during pheromone updates. (5) Smaller the value of $F(x)$ and $B(x)$ more amount of pheromone need to be updated. This also means the make that makespan and cost are smaller. Therefore the fitness is higher.

Step 7: The global resource manager maintain a table for each cloud .It contains the values of $F(x)$ and $B(x)$ and power (8)and energy consumption (9) objectives. Based on that the incoming request are analysed and scheduled to allocated resource.

VI. CONCLUSION

The Scheduling workloads in multi-cloud environment is challenging as the data centers have resources which are heterogeneous in nature and dynamic. This paper proposed an Improved ACO to solve Multi-objective Optimization problem. The task are scheduled efficiently to the allocated resource with less time and cost. It also avoid resource wastage and increase the resource utilization with reduced power consumption. So by using this both the user and provider will get satisfied.

REFERENCES

[1] R. Van den Bossche, K. Vanmechelen, and J. Broeckhove, ''Cost-efficient scheduling heuristics for deadline constrained workloads on hybrid clouds,'' in Proc. IEEE 3rd Int. Conf. Cloud Comput. Technol. Sci. (CloudCom), Nov./Dec. 2011, pp. 320–327.

[2] W. Daun, X. Fu, F. Wang, B. Wang, and H. Hu, ''Cloud computing task scheduling model based on improved ant colony algorithm,'' Comput. Eng., vol. 41, no. 2, pp. 12–16, Feb. 2015.

[3] Y. Zhou and X. Huang, ''Scheduling workflow in cloud computing based on ant colony optimization algorithm,'' in Proc. 6th Int. Conf. Bus. Intell. Financial Eng. (BIFE), Nov. 2013, pp. 57–61.

[4] Xia-yu Hua, Jun Zheng, Wen-xin Hu. Ant colony optimization algorithm for computing resource allocation based on cloud computing environment[J].Journal of East China Normal University(Natural Science), 2010,2(1):127-134.

[5] J. U. Duncombe, "Infrared navigation—Part I: An assessment of feasibility," IEEE Trans. Electron Devices, vol. ED-11, pp. 34-39, Jan. 1959.

[6] C. Y. Lin, M. Wu, J. A. Bloom, I. J. Cox, and M. Miller, "Rotation, scale, and translation resilient public watermarking for images," IEEE Trans. Image Process., vol. 10, no. 5, pp. 767-782, May 2001.

[7] S. K. Panda ; P. K. Jana, "A multi-objective task scheduling algorithm for heterogeneous multi-cloud environment" Electronic Design, Computer Networks & Automated Verification (EDCAV), 2015 International Conference on 29-30 Jan. 2015,pages 82-87

[8] M. Choudhary, Sateesh Kumar Peddoju, "A Dynamic Optimization Algorithm for Task Scheduling in Cloud Environment," vol. 2, no. 3, pp. 2564–2568, 2012

[9]G. Ming and H. Li, "An Improved Algorithm Based on Max-Min for Cloud Task Scheduling", Recent Advances in Computer Science and Information Engineering, Lecture Notes in Electrical Engineering, Vol. 125, pp. 217-223, 2012.

[10] K. Liu, H. Jin, J. Chen, X. Liu, D. Yuan and Y. Yang, "A Compromised-Time-Cost Scheduling Algorithm in SwinDeW-C for Instance-IntensiveCost-Constrained Workflows on a Cloud Computing Platform", The International Journal of High Performance Computing Applications, Vol. 24, No. 4, pp. 445-456, 2010.

[11] S. K. Panda and P. K. Jana, "An Efficient Task Scheduling Algorithm for Heterogeneous Multi-cloud Environment", 3rd IEEE International Conference on Advances in Computing, Communications and Informatics, pp. 1204-1209, 2014.

[12] T. D. Braun, H. J. Siegel, N. Beck, L. L. Boloni, M.Maheswaran, A. I.Reuther, J. P. Robertson, M. D. Theys, B. Yao, D. Hensgen and R. 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, Vol. 61, No. 6, pp. 810-837, 2001.

K.SOUNDARYA received B.E. Degree in Electronics and communication engineering from K.Ramakrishnan College of Technology, Trichy in 2014. She currently pursuing her Master degree in Mobile and pervasive computing at University College of Engineering (BIT Campus) ,Trichy.

Mrs .R. JAYAMALA, Asst. Professor under the Department of Information Technology at University College of Engineering (BIT Campus) Trichy. Her Research focuses on the cloud computing. She Published papers in 7 International Journals, 2 International Conference, and four National Conference.