Novel Approach of Energy aware task scheduling in Cloud environment with Ant colony optimization

Sheenu
PIET PATIALA (CSE DEPTT.)
Gurpreet Singh
A.P CSE Department CGC COE, LANDRAN
Er. Kavneet Kaur
Assistant Professor PIET Patiala

Abstract- Power consumption is one of the most critical problems in data centers. One effective way to reduce power consumption is to consolidate the hosting workloads and shut down physical machines which increase the carbon rate percentage, then reduce the consumption of energy by increasing the utilization. In this paper optimize the scheduling by Ant colony optimization, which increase the utility and reduce energy consumption and carbon print.

Keywords- Optimization, cloud, scheduling, ACO

I. INTRODUCTION

A cloud broker is a third-party individual or business that acts as an intermediary between the purchaser of a cloud computing service and the sellers of that service. In general, a broker is someone who acts as an intermediary between two or more parties during negotiations. The broker's role may simply be to save the purchaser time by researching services from different vendors and providing the customer with information about how to use cloud computing to support business goals. In such a scenario, the broker works with the customer to understand work processes, provisioning needs, budgeting and data management requirements. After the research has been completed, the broker presents the customer with a short list of recommended cloud providers and the customer contacts the vendor(s) of choice to arrange service. A cloud broker may also be granted the rights to negotiate contracts with cloud providers on behalf of the customer. In such a scenario, the broker is given the power to distribute services across multiple vendors in an effort to be as cost-effective as possible, in spite of any complexity that negotiations with multiple vendors might involve. The broker may provide the customer with an application program interface (API) and user interface (UI) that hides any complexity and allows the customer to work with their cloud services as if they were being purchased from a single vendor. This type of broker is sometimes referred to as a cloud aggregator.**Power usage effectiveness (PUE)** PUE is one of the most famous metrics used to measure energy efficiency of a cloud computing services and it is defined as the total power used by the data center (Pt) divided by the total power consumed by the ICT equipment (Ps), and it can defined by Equation 1 [6].

$$PUE=\frac{Pt}{Ps}$$

The ideal value for PUE is 1 if all of the power consumed by the servers account for all of the power delivered to the data center and it is impossible to have a PUE less than 1 [6]. Data from Uptime Institute surveys suggests that average PUE of today's data centers [6]. data center, which reported that the PUE values between 1.33 and 3, and the average value was 2.04. According to the report to congress on servers and data center energy efficiency, PUE ratio of 2.0 was assumed to be the average value across all U.S. data centers [24]. Levels of energy efficiency based on PUE are presented in [25]. A PUE value equal to or close to 1 is theoretically possible by spending zero energy on cooling. The near-zero cooling energy is possible by using free environmental cold-air-, water-, and evaporation-based cooling economizers such as in the Facebook data center [26].

Carbon usage effectiveness (CUE)

For such data centers which get their entire power from electric power distribution and do not leave the local carbon footprints, CUE is defined as equation 2 [27]. In contrast to PUE which is expressed as a number without unit, the units of the CUE metric are kilograms of carbon dioxide (kgCO2eq) per kilowatt-hour (kWh)

$$PUE = CEF \times PUE \quad (2)$$

II. LITERATURE REVIEW

Nikouiet. al. [1]In this paper, by considering the focal part of the Cloud Broker in distributed computing, "Two-Phase Carbon Aware Cloud Broker" has been suggested that endeavor to limit vitality and carbon by considering the vitality and carbon proficiency of server farms which may topographically be conveyed. For utilization up to 20%,15% in contrast with Round Robin deferentially.

Wadhwa, Bharti et.al.[2]this approach utilizes the carbon impression rate of the datacenters in conveyed cloud design and the idea of virtual machine assignment and movement for diminishing the carbon emanations and vitality utilization in the combined cloud framework. Reenactment comes about demonstrate that our proposed approach decreases the carbon dioxide discharges and vitality utilization of combined cloud server farms when contrasted with the established planning methodology of round-robin VM planning for unified cloud datacenters.

Capozzoli, Alfonso, et al. [3]In this paper a review of the principle imperative vitality and warm measurements is given. A basic analysis to research common relations among measurements was performed, with the plan to illuminate some physical viewpoints in regards to the evaluation of DC worldwide vitality execution. Undoubtedly, in spite of the fact that these measurements are normally used to survey the vitality effectiveness of DCs, their value for empowering lower vitality utilization was inadequately researched. In addition, an analysis of the impact of the DC warm execution on measurements was finished. The warm administration accept a key part in accomplishing vitality sparing amid the operation of a DC and for the change of the IT hardware unwavering quality.

Adamutheet. al. [4]In this paper, the virtual machine position issue is planned as a multi-target improvement issue. The targets are boosting benefit, amplifying load adjusting and limiting plan of action wastage. Aftereffects of Genetic calculations, Non-ruled Sorting Genetic Algorithm and Non-ruled Sorting Genetic Algorithm-II are contrasted and basic arrangement portrayal, punishment and advantage esteem. All the three calculations revealed great arrangements while GA and NSGA are subjected to untimely meeting and copy arrangements. NSGA-II gives a decent and differentiated scope of arrangements.

Malekloo et al. [5]this paper concentrates on this issue of virtual machine situation and proposes a multi-target advancement way to deal with limit both power utilization and asset wastage and to limit vitality correspondence cost between arrange components inside a server farm. An Ant Colony Optimization (ACO) calculation is proposed to get a Pareto set for a multi-target issue. The proposed calculations are tried utilizing Cloudsim instruments. The exhibitions of these calculations are contrasted and three surely understood single-objective methodologies and a multi-objective Genetic Algorithm (GA). The outcomes show that the proposed calculations can look for and discover arrangements that display adjust between various goals. In any case, ACO can And preferable arrangements over GA regarding our goals.

Mustafa, Saad, et al. [6] In this paper, they assess the execution of existing vitality effective BFD calculations in view of different workloads and movement methods. Additionally, considering the hugeness of Service Level Agreement (SLA), we present SLA-mindfulness in customary BFD calculation to limit the SLA infringement. We display the analysis and perceptions for each of the considered procedures in view of aggregate vitality utilization, normal SLA infringement, and SLA execution debasement because of movement.

Kar, Ipsitaet. al. [7]In this paper, they have a tendency to stipulate a structure that determines vitality minimization is a speculation of makespan limited by utilizing the Energy-Aware Task Scheduler utilizing Genetic Algorithm. We endeavored to direct a review of various booking techniques for vitality minimization in cloud server farms and their restrictions and toward the finish of this paper, we render a green advanced vitality mindful undertaking planning calculation for Cloud server farms.

Vakiliniaet. al.[8]The principle thought behind this paper is that with the joint effort of streamlining booking and estimation systems, the power utilization of DC can be ideally decreased. In the stage, an estimation module has been installed to anticipate the future heaps of the framework, and at that point, two schedulers are considered to plan the normal and unpredicted burdens, separately. The proposed scheduler applies the segment era method to handle the number straight/quadratic programming enhancement issue. Additionally, the cut-and-unravel based calculation and the get back to technique are proposed to lessen the many-sided quality and calculation time. At last, numerical and exploratory outcomes are introduced to approve our discoveries. Adjustment and versatility of the proposed stage result in a striking execution in VM position and movement forms. We trust that our work propels the cutting edge in workload estimation and dynamic power administration of cloud DCs, and the outcomes will be useful to cloud specialist organizations in accomplishing vitality sparing.

Abd, Sura Khalil, et al. [9]In this paper, they present a DNA-based Fuzzy Genetic Algorithm (DFGA) that utilizes DNA-based planning procedures to lessen control utilization in cloud data centers. It is a power-mindful design for overseeing power utilization in the distributed computing foundation. they likewise distinguish the exhibitions measurements that are expected to assess the proposed work execution. The test comes about demonstrate that DFGA diminished power utilization when contrasting and different calculations. Our proposed work manages constant undertaking which is not static, and focuses on the dynamic clients since they are associated with cloud.

Wilde, Torsten, et al. [10]To address these issues, this paper presents the measurements framework PUE (sPUE) and Data focus Workload Power Efficiency (DWPE). sPUE figures the overhead to operate a given framework in a specific server farm. DWPE is then computed by deciding the vitality proficiency of a particular workload and partitioning it by the sPUE. DWPE would then be able to be utilized to characterize the vitality proficiency of running a given workload on a particular HPC framework in a particular server farm and is right now the main completely coordinated metric appropriate for rating a HPC server farm's vitality effectiveness. What's more, DWPE takes into consideration foreseeing the vitality proficiency of various HPC frameworks in existing HPC server farms, subsequently making it a perfect approach for managing HPC framework acquirement. This paper closes with an exhibition of the utilization of DWPE utilizing an arrangement of delegate HPC workloads

III. PROPOSED METHODOLOGY

The Ant Colony Optimization (ACO) algorithm was introduced by Dorigo M. in 1996 based on the real ant behavior and it's a new heuristic algorithm to solve combinational optimization problems. Investigations have shown that ants have the ability to find food in an optimal path between the food and nest. With the ant motion some pheromone is released on ground, previous laid trail is encountered by isolated ant is being detected and follow with a higher probability. The ant's probability of choosing the way depends upon the pheromone concentration on that way. Higher the pheromone concentration, higher will be the probability of that way adoption. An optimal way can be found by utilizing this positive mechanism of feedback. The steps of Ant Colony Optimization algorithm are given below in Algorithm

Algorithm 1: ACO

Step 1: Parameters are set as an input; pheromone trails are initializing.

Step 2: On path segments, theVirtual trail is accumulated.

Step 3: ACO - Construct Ant Solutions

From node i to node j an ant will move with probability

$$P_{i,j}=\frac{(\tau_{i,j}^{\alpha})(\eta_{i,j}^{\beta})}{\sum(\tau_{i,j}^{\alpha})(\eta_{i,j}^{\beta})}$$

Where,

On edge i, j the amount of pheromone is $\tau_{i,j}$.To control the influence of $\tau_{i,j}$ α is a parameter

In edge i, j (typically $1/d_{i,j}$) $\eta_{i,j}$ is the desirability

To control the influence of $\eta_{i,j}$ β is a parameter

Step 4: ACO - Pheromone Update

According to the equation amount of pheromone is updated

$\tau_{i,j} = (1-\rho)\tau_{i,j} + \Delta\tau_{i,j}$

Where,

On a given edge i, j the amount of pheromone is $\tau_{i,j}$

ρ is the rate of pheromone evaporation is ρ

The amount of pheromone deposited is $\Delta\tau_{i,j}$, typically given by

$$\Delta\tau_{i,j}^{k} = \begin{cases} \frac{1}{L_k} & \text{if ant k travels on edge i, j} \\ 0 & \text{otherwise} \end{cases}$$

Where,

The cost of the k[th] ant's tour (typically length) is L_k

4. PROPOSED Scheduling ACO Algorithmto Calculate Carbon Foot Print Rate and Energy Consumption.

With the given Task Scheduling ACO algorithm (Algorithm.1) characteristic utilization, the task can be scheduled. Similarly, new task can be carried out with the utilization of previous task scheduling result. The basic ideas of ACO algorithm is inherited in WS-ACO algorithm for the reduction of execution time and cost.

Parameters

Carbon foot print rate = $c_i = \boldsymbol{p\_NUM_i \times p\_MIPS_i + VM\_b_i}$

Energy consumption= $\boldsymbol{\tau_i(0) = p_{NUM_i} \times p_{MIPS_i} + VM_{b_i} * Price\ per\ units}$

A. Initialize Pheromone

On VMs, there is a random distribution of ants at the beginning and then, VM_i pheromone values are initialized:

$$\boldsymbol{\tau_i(0) = p\_NUM_i \times p\_MIPS_i + VM\_b_i} \quad (1)$$

Where

p_NUM_i←number of VM_i processor

p_MIPS_i←million instructions per second of each VM_i processor

VM_b_i←VM_i communication bandwidth ability

B. Choosing VMs rule for next task

For next task, VM_i choose by k-ant with probability defined as:

$$\boldsymbol{P_i^K(T)} = \begin{cases} \frac{[\tau_i(T)]^{\alpha}[c_i]^{\beta}[lb]^{\gamma}}{\sum[\tau_K(T)]^{\alpha}[c_K]^{\beta}[lb]^{\gamma}} & \boldsymbol{if\ i \epsilon 1,2 \ldots\ldots\ldots.n} \\ \mathbf{0} & \boldsymbol{otherwise} \end{cases} \quad (2)$$

Where

$\tau_i(T)$←pheromone value of VM_i at time T

c_i←VM_i computing capacity

$C_i$can be defined as:

$$\boldsymbol{c_i = p\_NUM_i \times p\_MIPS_i + VM\_b_i} \quad (3)$$

lb_i←VM_i load balancing factor for minimizing the degree of imbalance defined as:

$$lb_i = 1 - \frac{et_i - Avg\_et}{et_i + Avg\_et} \quad (4)$$

Where

$Avg\_et \leftarrow$ virtual machine average execution time in the optimal path last iteration

$et_i \leftarrow$ expected execution time of VM_i task

et_i is defined as:

$$et_i = \frac{total\_TL}{c_i} + \frac{Input\_FS}{VM\_b_i} \quad (5)$$

Where

$total\_TL \leftarrow$ total length of task submitted to VM_i

$Input\_FS \leftarrow$ task length before execution

α, β and $\gamma \leftarrow$ parameters controlling the relative weight of pheromone trail along with VMs computing capacity and load balancing.

Once heavily loaded are some VMs becoming bottleneck in cloud influences the given task set makespan. The load balancing factor lb_i is defined in the ant algorithm for improving the capacity of lead balancing. Bigger the lb_i, higher will be the probability of choosing means VM_i comprehensive ability is greater now.

C. Updating Pheromone

Ant Let $\tau_i(T)$ at any time T be the VM_i pheromone intensity. The update of the pheromone is given by:

$$\tau_i(T+1) = (1-\rho) \times \tau_i(T) + \Delta\tau_i \quad (6)$$

Where

$\rho \epsilon [0,1] \leftarrow$ decay coefficient of pheromone trail

The past solution impact will be less if value of ρ is greater. The $\Delta\tau_i$ value is defined as:

After the completion of ant tour, updating the local pheromone on VM visited and $\Delta\tau_i$ value is given as:

$$\Delta\tau_i = 1/t_{iK} \quad (7)$$

Where

$t_{iK} \leftarrow$ K-ant searched shortest path length at i^{th} iteration

In case, the current optimal solution is found by the ant while completing its tour, larger intensity pheromone is laid on its tour and updating the global pheromone on VM visited and $\Delta\tau_i$ value is given as:

$$\Delta\tau_i = d/t_{op} \quad (8)$$

Where

$t_{op} \leftarrow$ current optimal solution

d $\leftarrow$ encouragement coefficient

SYSTEM MODEL

System model is a conceptual representation of working process of the proposed system. In system, model generally four basic elements are used to represent the interaction of different process with each other. When any model is proposed, we have to describe the internal flow of information from one process to another. With the help of system, we can show what types of processing taking place and then what type of information is going to another process as shown in Figure 1.

The proposed system model achieving the goal of optimized cost and time through following steps:

Step 1: Input the Task to the task simulator

Step 2: Parse the task from this job.

Step 3: Task follows the Differential Parse Distribution.

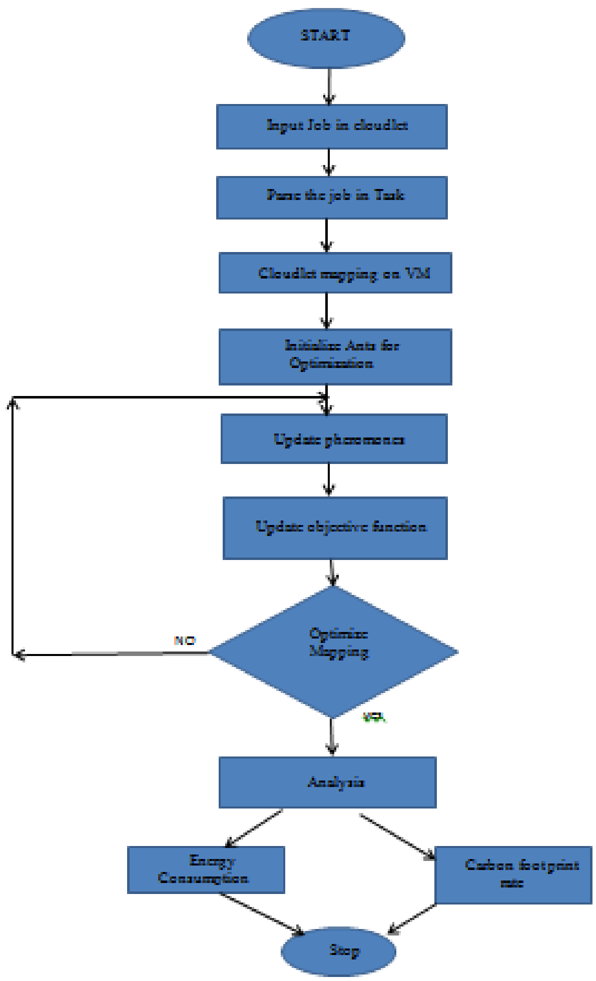
Step 4: According to the critical path, Virtual machine is provided to the task.

Step 5: Initialize the optimize parameter.

Step 6: By fitness function update the optimize parameter.

Step 7: Objective function is optimized by fitness function values.

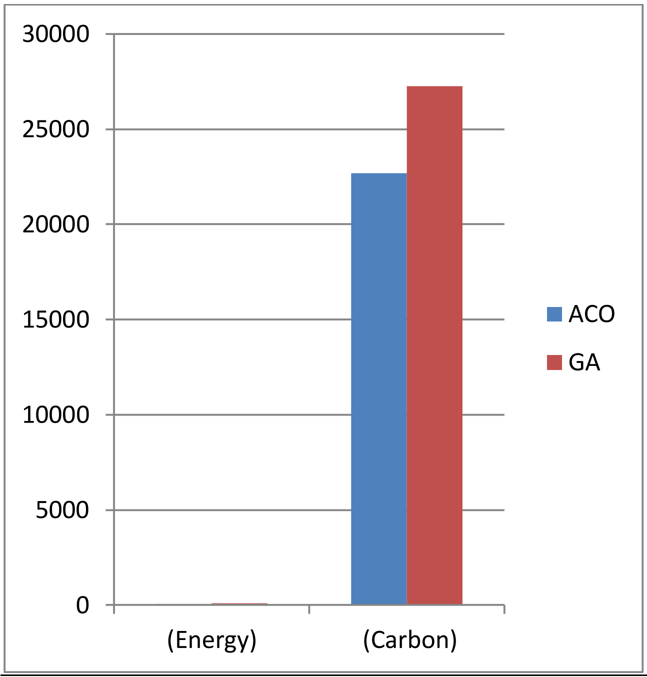
Step 8: If the function is optimized then we analysis the cost and time of that function. If the function is not optimized, then retransmit to step 5



IV.RESULTS

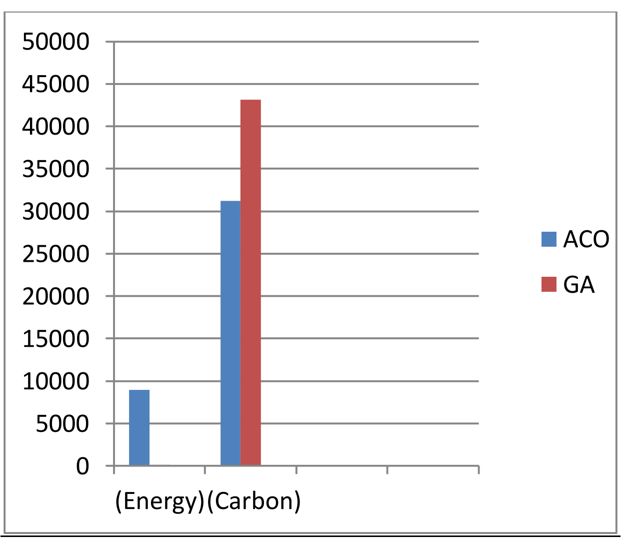
For 2

Parameters	ACO	GA
(Energy)	59.81	85.22
(Carbon)	22671.23	27262.64



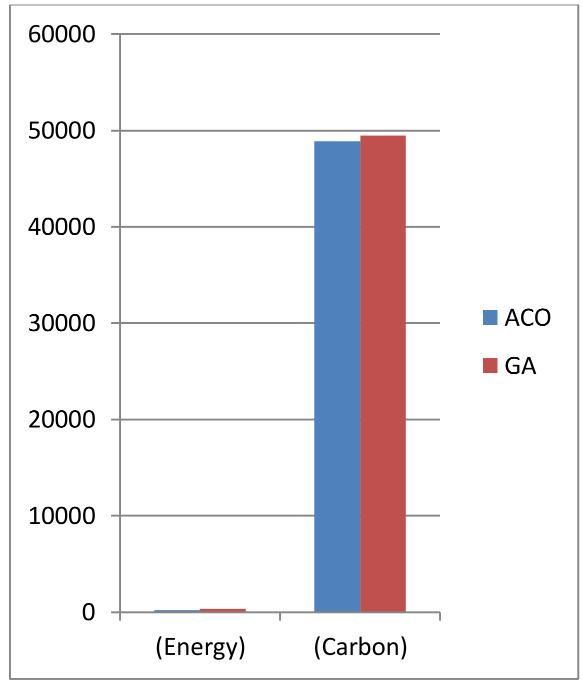
For 4

Parameters	ACO	GA
(Energy)	8993	112.63
(Carbon)	31229.51	43150.51



For 6

Parameters	ACO	GA
(Energy)	206.34	303.92
(Carbon)	48903.61	49449.94



V. REFERENCES

[1]. Nikoui, Tina Samizadeh, Sam Jabbehdari, and AlirezaBagheri. "Providing a Cloud Broker-based Approach to Improve the Energy Consumption and Achieve a Green Cloud Computing." *International Journal of Computer Applications* 138 (2016): 42-49.

[2]. Wadhwa, Bharti, and AmandeepVerma. "Carbon efficient VM placement and migration technique for green federated cloud datacenters." *Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on*. IEEE, 2014.

[3]. Capozzoli, Alfonso, et al. "Review on performance metrics for energy efficiency in data center: The role of thermal management." *International Workshop on Energy Efficient Data Centers*.Springer, Cham, 2014.

[4]. Adamuthe, Amol C., Rupali M. Pandharpatte, and Gopakumaran T. Thampi. "Multiobjective virtual machine placement in cloud environment." *Cloud & Ubiquitous Computing & Emerging Technologies (CUBE), 2013 International Conference on*.IEEE, 2013.

[5]. Malekloo, Mohammadhossein, and Nadjia Kara. "Multi-objective ACO virtual machine placement in cloud computing

environments." *Globecom Workshops (GC Wkshps), 2014*.IEEE, 2014.

[6]. Mustafa, Saad, et al. "Performance Evaluation of Energy-aware Best Fit Decreasing Algorithms for Cloud Environments." *Data Science and Data Intensive Systems (DSDIS), 2015 IEEE International Conference on*.IEEE, 2015.

[7]. Kar, Ipsita, RN RamakantParida, and Himansu Das. "Energy aware scheduling using genetic algorithm in cloud data centers." *Electrical, Electronics, and Optimization Techniques (ICEEOT), International Conference on*.IEEE, 2016.

[8]. Vakilinia, Shahin, BehdadHeidarpour, and Mohamed Cheriet. "Energy Efficient Resource Allocation in Cloud Computing Environments." *IEEE Access* 4 (2016): 8544-8557.

[9]. Abd, Sura Khalil, et al. "An effective approach for managing power consumption in cloud computing infrastructure." *Journal of Computational Science* 21 (2017): 349-360.

[10]. Wilde, Torsten, et al. "DWPE, a new data center energy-efficiency metric bridging the gap between infrastructure and workload." *High Performance Computing & Simulation (HPCS), 2014 International Conference on*.IEEE, 2014.