Task Scheduling In Cloud Platforms Using The Bat-PSO Hybrid Approach

Ms. Bhawana Lakhani[1], Prof. Amit Agrawal[2]
[1]Dept of Computer Science
[2]Professor, Dept of Computer Science
[1, 2] Medicaps University, Indore (M.P.)

Abstract- *Task Scheduling is the process of assigning tasks to different virtual machines over cloud based servers with an aim to optimize the resource utilization for cloud based platforms. Task scheduling is generally a complex task often leading to the solution to an optimization problem. In this paper, a hybrid approach employing the bat and particle swarm optimization (PSO) has been proposed. It has been shown that the proposed approach achieves better results compared to the hybrid of bat and heuristic approach. The performance evaluation parameter has been chosen to be the response time. The proposed system has been simulated on simulated on the CloudSim 3.0. with java in Netbeans.*

Keywords- Resource utilization, task scheduling, bat, particle swarm optimization (PSO), response time.

I. INTRODUCTION

Cloud computing allows several applications and services which would not be feasible to implement on normal computing machines. A cloud platform can be assumed to be a remote machine which can be accessed from a machine of the user by creating a virtual machine to emulate the interface of the actual cloud server. Cloud based platforms provide several applications such as web services, security services, big data and machine learning services, e-commerce based services to name a few. It aims at accessing a virtual machine that emulates the actual cloud server on a remote client machine. Several application today require cloud based services such as big data analytics, artificial intelligence based GPU access, data warehousing etc. Resource scheduling of cloud based platforms is necessary to adjust or to share the load evenly often termed as load balancing. Several approaches have been put forth thus far with their own merits and limitations. The following figure depicts the cloud deployment model for cloud computing applications.

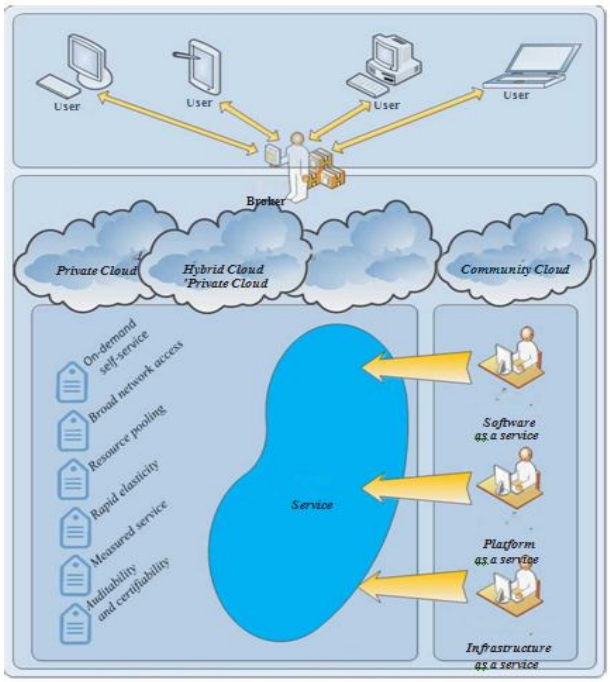

Fig.1 Deployment model for Cloud Computing

The figure above shows the cloud based services along with the deployment architecture.Cloud Computing holds the strong potential of providing scalable, flexible and economical solutions for the meeting the increasing storage and service requirements. Today it is being counted as one of the high end and demanding technologies. There is no second thought to the fact that cloud services provides plethora of benefits in many form of environments. But to use the services of cloud efficiently, the system administrators require the capability to optimally implement task scheduling.

II. RELATED WORK

This section highlights prominent work in the domain with their salient features.

F Juarez et al. [1] proposed an energy aware scheduling for task parallel task based applications in cloud computing. The aim was to make the scheduling scheme low energy consuming with a holistic aim of reducing the carbon footprint

of cloud based applications. It used an energy aware run time scheduler which cut short the energy consumption. The algorithm lies in eta heuristic resource allocation for implementation of the polynomial time algorithm.

Lan Wang et al. [2] provided Adaptive Dispatching of Tasks in Cloud. Cloud computing domain has been witnessing a large traffic and users dependent on it. With most of the work being shifted to the internet platform, the cloud services have become dominant in all aspects of business and technology. In this work, the authors proposed a novel study of cloud tasks dispatching. There are allocation schemed and algorithms that have been used as a part of the model. The response time that is computed has been reduced considerably in this work. Various hosts have been deployed for the proper client and server interaction. The time delays were greatly lessened and it proved to be a really useful methodology.

Martin Duggan et al. [3] presented Research on Predicting Host CPU Utilization in Cloud Computing using Recurrent Neural Networks. This study aims to predict the CPU consumption of host machines by using recurrent neural networks. The process involved utilizing the recurrent neural networks that could accurately predict the time series data and also collect the information with flexibility. With respect to the traditional approaches and methods, this method was successful in accurate forecasting and gave better outcomes.

Ning Liu et al. [4] propose A Hierarchical Framework of Cloud Resource Allocation and Power Management Using Deep Reinforcement Learning. It stood for a novel hierarchical framework that could address and solve all the possible power and resource allocation problems in the cloud based platforms. The proposed system took into account the virtual machines servers and various resources. The rising use of the reinforced deep learning solutions also helped in restructuring the entire concept and model.

Liyun Zuo et al. [5] proposed A Multiqueue Interlacing Peak Scheduling Method Based on Tasks Classification in Cloud Computing. It was mainly a scheduling scheme that was further improved. The resource allocation and tasks classification was carried out on the basis of the type of memory and the CPU consumption. The infrastructure within the workloads may vary. Put together, they give rise to a complete cloud solution. The CPU specific tasks were classified differently and the peak scheduling was used for it.

Yazhou Hu et al. [6] propose Three Models to Predict the Workload Based on Analysing Monitoring Data. The dataset for the cloud workload is a very important part of gauging the entire system design. The authors proposed three models for forecasting the cloud workload. And the help was taken from the dataset for the workload. By monitoring the data and information flow, it is easy to predict the workload extent and its quantity. This helps in building elasticity and also enhances the scalability of the system. The workload plays a crucial role and it must be flexible enough so that different programs can use it according to its changing requirements.

Ji Xue et al. [7] put forth PRACTISE, a neural network based framework that could predict the future cloud workloads, peak loads etc. The cloud workload prediction has been a very active area of research and the authors primarily focused on forecasting the peak loads and their timings etc. As due to overflow of data and resources, the cloud serves hold the probability to crash and go off. So, forecasting helps in giving optimization solutions to the problems faced.

Mehmet Demirci [8] gave a Survey of Machine Learning Applications for Energy-Efficient Resource Management in Cloud Computing Environments. The resource management and resource based allocation is very important area of work. With the advent of machine learning approaches and mechanisms, it is viable option to use them for the energy efficient resource handling for cloud based environments. There is no second thought to the fact that cloud services provides plethora of benefits in many form of environments.

Sherif Abdelwahab et al [9] proposed a technique for enabling Smart Cloud Services through Remote Sensing: An Internet of Everything Enabler survey. The concept of remote sensing has been utilized in this research work. The use of the cloud services by the IoT and remote sensing techniques has been the subject of the study. It is a very good concept to use both the technologies together. Already the emergence of the IoT has been coupled with the cloud based services and their amalgamation has been quite a success. This survey points out the aspects of the remote sensing for cloud services and other allied areas where it can be applied. Cloud based platforms provide several applications such as web services, security services, big data and machine learning services.

Chin-Feng Lai et al [10] proposed a Co-Processing Intermediary Framework Integrated Cloud and Wireless Body Sensor Networks. It is mainly based on the wireless sensor networks that are used for the cloud based approaches. It is mainly applied in the fall detection schemes. Using this methodology, the transmission time and data sensing time are lessened considerably and gives impressive performance. Today it is being counted as one of the high end and demanding technologies. There is no second thought to the fact that cloud services provides plethora of benefits in many form of environments.

III. PROPOSED SYSTEM

The designed system proposes a hybrid of the bat an particle swarm optimization techniques for task scheduling.

3.1 The bat optimization

The bat algorithm is often used for multi-objective optimization problems. In case of the cloud based approach, there are several objectives of an optimization technique namely reducing response time, enhancing throughput etc. It is often used for complex optimization problems with challenging constraints. The bat algorithm is based on the echolocation characteristics of microbats. The most important attribute is the updating rule for the positions and velocities for the micro bats in 'd' dimensional search space.

Let the position and velocities be designated by x_i^t and v_i^t
Here,

t represents the time steps
i represents the iteration
For the time interval $(t+1)$,

$$v_i^{t+1} = v_i^t + (x_i^t - x_*)f_i \quad (1)$$

and

$$x_i^{t+1} = x_i^t + v_i^t \quad (2)$$

here,

x_* is the current global best location which is computed after comparing all the locations for 'n' bats at iteration 't'.
f_i is the frequency of sending out the sampling values
.
The algorithm for implementing the multi-objective bat algorithm for task scheduling can be given by:

Step.1 Define objective functions

$$f1(x), f2(x) \ldots . fk(x) \quad (3)$$

Here,

$$x = (x_1 \ldots\ldots\ldots\ldots x_d)^T \quad (4)$$

Step.2 Initialize bat population with n=1

for (j=1 to N),

generate weights (K) such that

$$w_K \geq 0 \quad (5)$$

And

$$\sum_{k=1}^{K} w_k = 1 \quad (6)$$

Step.3 For a single objective function,

$$f = \sum_{k=1}^{K} w_k f_k \quad (7)$$

Step.4 while (t<max iterations),

Generate updated solutions

if$(rand > r_i)$
Implement random walk in the periphery of the best solution
else
Generate new solution by flying randomly

Step.5 Increment r_i and accept new solutions

Step.6 Rank the bats and find current best location x_*

Step.7 Record x_* as the non dominated solution

The next step is to further fine tune the results using the particle swarm optimization given by:

$$v_{id}(t) = w \times v_{id}(t-1) + c_1 \emptyset_1 (p_{id} - x_{id}(t-1)) + c_2 \emptyset_2 (p_{gd} - x_{id}(t-1)) \quad (8)$$

$$x_{id}(t) = x_{id}(t-1) + v_{id}(t) \quad (9)$$

The parameters used are:

v	The particle velocity
x	The particle position
t	Time
c_1, c_2	Learning factors
Φ_1, Φ_2	Random numbers between 0 and 1
p_{id}	Particle's best position
p_{gd}	Global best position
w	Inertia weight

IV. SIMULATION AND RESULTS

The system is simulated on the CloudSim 3.0. with java in Netbeans.

Entity	Parameter	Value
Task (as Cloudlet)	task_length	1000-40000
	Total_no._of_task	10-200
Virtual Machine(VM)	Total no_of_VMs	8-32
	MilliInstPerSec,(mips)	500-2000
	VM_memory(RAM)	512
Data-center	No_of_Data_center	10
	No_of_Host	2-6

Fig.2Parameter Specifications

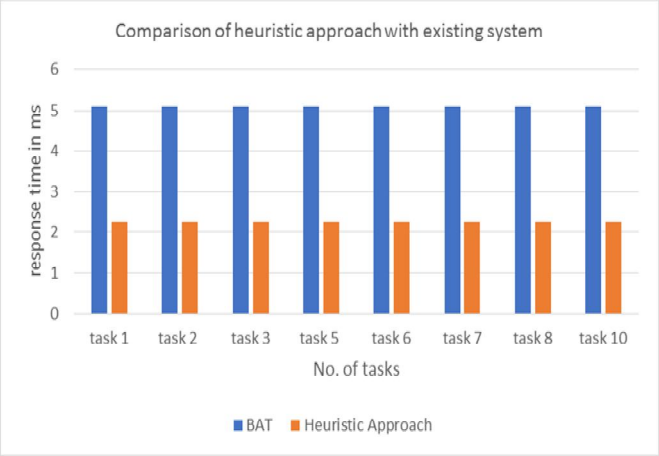

Fig.3 Comparison of response time for bat and heuristic approaches

V. CONCLUSION

It can be concluded from previous discussions that the proposed system attains lower response time (in ms) with the hybrid bat & PSO technique compared to the previously existing bat & heuristic approach. The effect of the results would mean a faster scheduling implantation for load balancing of the system making it well suited for big data applications.

REFERENCES

[1] F Juarez, J Ejarque, RM Badia, "Dynamic energy-aware scheduling for parallel task-based application in cloud computing", 2018 ELSEVIER.

[2] Lan Wang and Erol Gelenbe, "Adaptive Dispatching of Tasks in the Cloud", 2018 IEEE.

[3] Martin Duggan, Karl Mason, Jim Duggan, Enda Howley, Enda Barrett, "Predicting Host CPU Utilization in Cloud Computing using Recurrent Neural Networks", 2017 IEEE.

[4] Ning Liu, Zhe Li, Jielong Xu, Zhiyuan Xu, Sheng Lin, Qinru Qiu, Jian Tang, Yanzhi Wang, "A Hierarchical Framework of Cloud Resource Allocation and Power Management Using Deep Reinforcement Learning", 2017 IEEE.

[5] Liyun Zuo, Shoubin Dong, Lei Shu, Senior Member, IEEE, Chunsheng Zhu, Student Member, IEEE, and Guangjie Han, Member, IEEE, "A Multiqueue Interlacing Peak Scheduling Method Based on Tasks' Classification in Cloud Computing", 2016 IEEE.

[6] Yazhou Hu, Bo Deng, Fuyang Peng and Dongxia Wang, "Workload Prediction for Cloud Computing Elasticity Mechanism", 2016 IEEE.

[7] Ji Xue, Feng Yan, Robert Birke, Lydia Y. Chen, Thomas Scherer, and Evgenia Smirni, "PRACTISE: Robust Prediction of Data Center Time Series", 2015 IEEE.

[8] Mehmet Demirci, "A Survey of Machine Learning Applications for Energy-Efficient Resource Management in Cloud Computing Environments", 2015 IEEE.

[9] Sherif Abdelwahab, Member, IEEE, Bechir Hamdaoui, Senior Member, IEEE, Mohsen Guizani, Fellow, IEEE, and Ammar Rayes, "Enabling Smart Cloud Services Through Remote Sensing: An Internet of Everything Enabler", 2014 IEEE.

[10] Chin-Feng Lai, Member, IEEE, Min Chen, Senior Member, IEEE, Jeng-Shyang Pan, Chan-Hyun Youn, Member, IEEE, and Han-Chieh Chao, Senior Member, IEEE, "A Collaborative Computing Framework of Cloud Network and WBSN Applied to Fall Detection and 3-D Motion Reconstruction", 2014 IEEE.

[11] Ian Davis, Hadi Hemmati, Ric Holt, Mike Godfrey, Douglas Neuse, Serge Mankovskii, "Storm Prediction in a Cloud", 2013 IEEE.

[12] Abul Bashar, "Autonomic Scaling of Cloud Computing Resources using BN-based Prediction Models", 2013 IEEE.

[13] Sadeka Islam , Jacky Keunga, Kevin Lee, Anna Liu, "Autonomic Scaling of Cloud Computing Resources using BN-based Prediction Models", 2012 ELSEVIER.

[14] Erol Gelenbe, Ricardo Lent and Markos Douratsos, "Choosing a Local or Remote Cloud", 2012 IEEE.

[15] Mohammad Moein Taheri and Kamran Zamanifar, "2-Phase Optimization Method for Energy Aware Scheduling of Virtual Machines in Cloud Data Centers", 2011 IEEE.

[16] Saurabh Kumar Garg, Srinivasa K. Gopalaiyengar, and Rajkumar Buyya, "SLA-Based Resource Provisioning for Heterogeneous Workloads in a Virtualized Cloud Datacenter", 2011 IEEE.

[17] Md. Toukir Imamt, Sheikh Faisal Miskhatt, Rashedur M Rahmant, M. Ashraful Amin, "Neural Network and Regression Based Processor Load Prediction for Efficient Scaling of Grid and Cloud Resources", 2011 IEEE.
[18] Christopher Stewart, Kai Shen, Arun Iyengar, Jian Yin, "EntomoModel: Understanding and Avoiding Performance Anomaly Manifestations", 2010 IEEE.
[19] Maria Couceiro, Paolo Romano, Lu´ıs Rodrigues, "A Machine Learning Approach to Performance Prediction of Total Order Broadcast Protocols", 2010 IEEE.
[20] Haibo Mi, Huaimin Wang, Gang Yin, Yangfan Zhou, Dianxi Shi, Lin Yuan, "Online Self-reconfiguration with Performance Guarantee for Energy-efficient Large-scale Cloud Computing Data Centers", 2010 IEEE.
[21] H Zhang, G Jiang, K Yoshihira, H Chen, "Intelligent workload factoring for a hybrid cloud computing model", IEEE 2009
[22] H Zhang, G Jiang, K Yoshihira, H Chen, "Resilient workload manager: taming bursty workload of scaling internet applications", IEEE 2009
[23] A Quiroz, H Kim, M Parashar, "Towards autonomic workload provisioning for enterprise grids and clouds", IEEE 2009
[24] A Stage, T Setzer, "Network-aware migration control and scheduling of differentiated virtual machine workloads", ACM 2009
[25] C Moretti, K Steinhaeuser, D Thain, "Scaling up classifiers to cloud computers", IEEE 2008