Comparative Analysis of Task Scheduling Algorithms for Energy Utilization in Green Cloud Computing using Soft Computing Techniques.

Ms. Veena M[1], Ms Priyanka V[2]
[1]*Assistant Professor*, [2]*MCA Student*
Department of Computer Applications, Presidency College, Bengaluru, India.

Abstract **-** Internet has become an important component in today's life. Huge amount of data is transferred and stored. Organization have moved towards Cloud computing and it is increasing everyday in a tremendous amount. There is also an increase in demand for services. Cloud computing is offering the services in terms of storage and software on demand. Data is stored in remote servers which are data centers. Number of users using cloud is increasing everyday as the services provided by the cloud is excellent. Due to this increase of computing, the energy used in datacenters needs more attention in case of the amount of energy consumed. There is a need for the management in Technology section as it is creating harmful impact to our environment because there is an increase in the carbon foot print. This needs to shift to green cloud computing. We have proposed optimal task scheduling using Artificial neural network and meta heuristic algorithms for the efficient utilization of energy. A comparative analysis of various task scheduling algorithms using soft computing techniques are given and a model is proposed for scheduling the task.

Keywords **-** cloud computing, green cloud computing, datacenters, energy consumption, Artificial neural networks.

I. INTRODUCTION

The rapid development of processing, storage technologies and increase in the speed of the internet has enabled cloud computing. As per National institute of standard and technology "Cloud computing is a model for on demand services by sharing various computing resources[2]. It gives power to access data, storage, applications through the internet in a convenient way. The network services and computing services are given because of the availability of high speed Internet. Providing these network services to storage and computing is the main purpose. The surplus advantages of cloud computing gives the customers an edge over the traditional computing.

Rather than having a dedicated infrastructure, the services can be rented on demand in cloud and cost effective manner. As a result of this most of the industries have shifted to cloud computing.

A. The cloud deployment models –

A Private cloud: A Cloud service available only to one particular organization. It is limited with the participants to only pertained that organization. It is only restricted to few people and it is more secured but expensive.

A Public Cloud: A cloud which is open generally for the public. It is managed and maintained by the academics, government organization, business and some combinations of them. Data is shared among the users and it is less secured. Some of the public clouds are Microsoft Azure, google, Allibaba, etc

B. Service Models of cloud Computing -

Infrastructure as a service(Iaas): provides on demand infrastructural service in terms of storage, servers, data centers network resources. The user need not have to purchase but can rent and pay only for the duration used. Examples include Amazon EC2, IBM SmartCloud Enterprise.

Platform as a service(Paas): provides all the applications, operating system and database deployed by the cloud users. Examples include Microsoft Azure, Google compute Engine etc.

Software as a service(Saas): Provides on demand cloud applications with automatic scaling capacity. Examples include Facebook, youtube, Google Apps etc.

In Figure1,various models of cloud computing with examples are given.

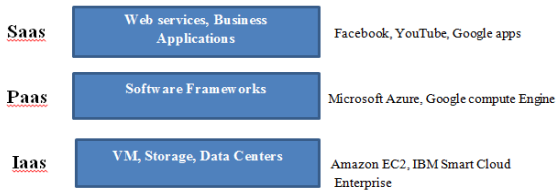

Figure 1: Service models of cloud computing

C. Green Cloud computing -

Green cloud is a model that helps in reducing the energy consumption based on few techniques. Some of the techniques are VM consolidation, load balancing, Virtualization, Relocation in virtual machines, merging. Task scheduling is one among them. Where in this technique will make use of efficient task scheduling algorithms are used to allocate the task to virtual machines so that all the tasks are completed in less time. It in turn optimizes energy consumption. The Green cloud computing which is useful in lessening of Green house gases which will consequently supportive in diminishing the worldwide cautioning.

Green cloud computing architecture: It basically includes three tiers that is first top most tier is for customer where customer can make request for services provided by cloud and customer is free to select the cloud from which services are required.

Second level or the middle level includes the application where it acts like cloud interface for customers and it also define the cloud application requirement.

Third level includes the servers where it is real cloud in which actual execution and storage of data happens

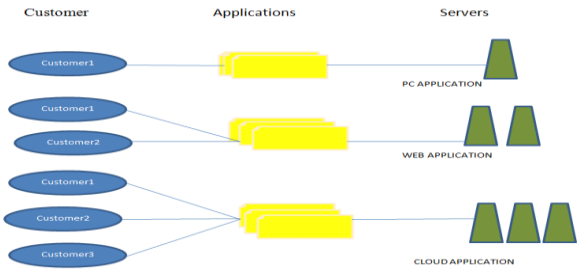

Figure 2: Green cloud computing architecture

D. Data Centers -

It is the heart of cloud computing and critical part of an organization. Provides facility in one place where many servers or machines are engaged in collecting, storing, processing and distribution of massive amounts of data. Large amount of data is getting created as people are everyday doing social networking, using Business applications, etc. [18]. Data centers are usually huge building where the machines are running continuously. There is a need of more energy in order to run all the components. In a survey was found up to 3% of all US electricity is consumed by data centers[19]. The amount of energy consumed by the datacenters is doubled every four. Data centers need around 200TWh of energy that is more than the national energy consumption of some of the countries. Natural Resources Defense council has mentioned that datacenters use 91billion kilowatts of energy by 2013 and they will use 139 billion kilowatts of energy by 2020 and contributing to 200 million metric tons of carbon di oxide[8]. They are contributing up to 0.3% of carbon emission.

Optimizing energy consumption is essential as data centers release large amount of heat which is not eco friendly to the environment as there is more release of co_2 adding to global warming. Whenever servers are not in use they can switch to stand by mode. Servers which are efficient use less power. A better techniques are needed to optimize the energy efficiency using optimal energy efficiency tools[20].

E. Measurement of Energy Consumption-

A several energy consumption benchmarks are used to measure power consumption at datacenters. These parameters helps to decide the efficient usage of the power and eco friendly data centers. The five important parameters used are[22]:

Power usage effectiveness (PUE): It was introduced in 2006 for Green Grid and has become global standard in 2016 by ISO/IEC9[16].It is the ratio of the total amount of energy consumed by the datacenter facilities to the energy being consumed by the IT infrastructure. PUE 1.0 means total energy is consumed and efficiently. PUE 2.0 means some energy is spent for cooling.

PUE = (Total data center power) / (Total IT infrastructure power)

Datacenter energy productivity (DCEP): It was introduced in 2007. DCEP measures the efficiency along with the amount of productivity. It is the ratio of the total useful work with the amount of total energy consumed.

DCEP = (Total useful work produced) / (Total energy consumed by datacenter)

Energy Reuse Factor (ERF): It is the ratio of the total data center energy reused elsewhere with the total data center energy. ERP will range from 0 to 1.0. ERP 0 means no energy is reused, and ERP 1.0 means all the energy is reused.

ERP = (Total energy reused) / (Total energy of data center)

Green Energy Coefficient (GEC): It is the measure of amount of renewable energy used by the data centers. It is the ration of the total amount of renewable energy used with the total amount energy consumed by the data center. GEC with 1.0 means all the energy used is renewable energy.

GEC = (Total renewable energy consumed) / (Total energy consumed)

Carbon usage effectiveness (CUE): It is the measure of the total amount of various gases emitted carbon di oxide etc. It is the ratio of the amount of carbon di oxide emitted with the total energy consumed. CUE 0.0 means no carbon di oxide emitted and it is an ideal value.

CUE = (Total amount of carbon di oxide emitted) / (Total energy consumed by data center)

F. Soft Computing -

There are many real time problems which is difficult to solve practically as it requires more computational time. One such problem is the task scheduling as it comes under NP hard problem[15]. Some times these problems can be solved by biologically inspired techniques which are more efficient.

Soft computing was proposed by Lotfi A Zadeh are nature inspired based solving the problems. It is different from the conventional technique as it is tolerant for imprecision, partial truth, uncertainty and approximation. Soft computing is just the automating the process of computing. It consist of distinct techniques, concepts which aim to solve the difficulties encountered in real world problems. In this the programs learns on its own from previous information. Basically soft computing techniques are Fuzzy logic, Evolutionary algorithms, Neural Networks.

G. Artifical Neural Network(ANN) -

It is a soft computing technique. ANN is an information processing which are built by interconnecting elementary processing systems called neurons. ANN artificially

represents human brain, which tries to simulates learning process. It is an interconnection of artificial neurons that uses a mathematical model or computational model for the learning process. ANN maps the input data with the approximate output. It consists of input layer which receives input, output layer that provides output and between these a hidden layer. ANN has a broad applications clustering, pattern recognition, function approximation and predicting systems.

Task Scheduling in data centers is multiprocessor task scheduling. This can be done by heuristic and meta heuristic approach. In heuristic approach a queue is maintained for the tasks. This queue is a priority queue and tasks are scheduled in first come and first serve bases. In meta heuristic a frame work is given with some general directions or strategies on how to solve a problem. They are heuristic but they work efficiently.

Several algorithms are developed using evolutionary algorithms of soft computing. A meta heuristic algorithms can be used to train ANN model for scheduling the task efficiently with minimum completion time to attain a better utilization of energy.

II. LITERATURE SURVEY

In Green Cloud Computing to reduce carbon footprint efficient utilization of energy is required. The energy consumption in data centers are done to hardware based and software based. Some of the techniques applied are virtualization, relocation in virtual machines (VM), merging, load balancing, VM consolidation, task scheduling. Scheduling the task to the resources is important for the efficient utilization of the energy. Basically task scheduling is the optimization problem and many soft computing techniques are used. Optimally scheduling the task in servers reduces the energy consumption. Many researchers have proposed different algorithms in the area of task scheduling using soft computing techniques.

Gan Guo-ning et al[1], developed a optimized task scheduling algorithm using genetic simulated annealing for cloud computing . Quality of service parameters for the different task in different aspect are given. The algorithm efficiently completes the scheduling of the task.

Linan Zhu et al[3], developed a resource scheduling in cloud computing using Ant colony Optimization. Experiment was done and the results showed that the algorithm had good performance on load balancing.

Pradeep Kumar et al [4], proposed a new scheduling algorithm "Improved Genetic algorithm". In this algorithm Min-Min and Max-Min scheduling techniques are combined to efficiently schedule the task in multiple machines.

Jianhua Gu et al [5],developed a scheduling algorithm on load balancing for Virtual Machines using Genetic Algorithm. The Algorithm had good efficiency and was able to solve load imbalance and high migration after Virtual machine scheduled.

Medhat A.Tawfeek et al[6],developed a task scheduling algorithm in cloud computing, using Ant colony optimization technique. In this algorithm makespan of the task was minimized.

Jing Liu et al[7],developed task scheduling model applying multi objective genetic algorithm (MO-GA) to minimize energy consumption and the algorithm gives better solution.

Liji Jacob et al[9],developed resource scheduling algorithm in Cloud Computing by using Bacterial for aging optimization algorithm. The was able to reduce, Cost, Makespan and improved reliability.

Wanneng shu et al[10],developed resource allocation algorithm based on clonal optimization, which improves energy efficiency and response time and makespan.

A.I Awad et al[11],Proposed model for scheduling and allocation using load balancing mutation particle swarm optimization. This model showed that it can save in makespan, execution time, transmission cost and round trip time.

Mohammed Shojafer et al[12],Proposed an algorithm using Fuzzy Theory and genetic algorithm. The algorithm did optimal load balancing by considering execution time and cost.

R K Jena [13], developed a task scheduling algorithm using multi objective nested particle swarm optimization which optimizes energy and processing time. The results showed that maximum scheduling is done in less time.

M.Kowsigan et al[14],developed an scheduling and load balancing algorithm combining Fussy set and Genetic Algorithm. It reduced makespan.

Ipsita Kar et al[17],Developed Energy aware task scheduling using Genetic Algorithm. Energy minimization is done by minimizing makespan.

Huda Ibrahim et al[21], developed a task scheduling algorithm using Integer Linear Programming which minimizes energy consumption in cloud. Later an adaptive Genetic Algorithm which minimizes energy consumption in dynamic cloud and optimally schedules task for minimizing the energy consumption.

A Comparative Study of various scheduling techniques is given in the Table 1

Tabel 1: Comparative study of task scheduling techniques

Author	Techniques used	Benefits	Scheduling Techniques
Gan Guo [1]	Based on Genetic simulated annealing algorithm.	Algorithm Efficiently schedules the task using parameters in various dimensions.	Implemented by combing Genetic Algorithm(GA) and Simulated Annealing
Linanzhu [13]	Based on Ant Colony Optimization (ACO)	Takes less time compared with the general algorithm	ACO and Random distribution algorithm.
Pradeep Kumar [4]	Based on Min-Min, Max-Min and genetic algorithm.	Tasks completed in minimum time and in efficient manner.	Implemented by combining Min-Min algorithm, Max-Min algorithm and Genetic algorithm.

Jianhua Gu [5]	Load balancing and scheduling strategy using GA for VM resource	Efficiently schedule the load balance for VM resources.	Using GA
Medhat A Tawfeek et al [6]	Based on ACO technique	Efficient task scheduling and minimizing makespan	Based on ACO and a simple heuristic is used.
Jing Liu [7]	Based on multiobjective genetic algorithm(MO-GA)	A better task scheduling to minimize energy consumption	Pareto Solution in MO-GA
Liji Jacob [9]	Based on Bacterial foraging optimization algorithm(BFOA)	Minimizes makespan, reduces cost	Applied hyper heuristic for resource scheduling
Wanneng shu [10]	Improved clonal selection algorithm (ICSA)	Improves energy consumption,makespan and reliablility	ICSA
A.I Awad[11]	Using load balancing PSO	Improves on makespan, execution, round trip,transmission cost.	Each task allocated to VM and PSO selects optimal distribution of task.
M.Shojafar et al[12]	Fuzzy theory and GA	Improves performance on makespan,, execution cost by 45% and total execution time by 50%	Fuzzy theory is used to get values for the fitness of chromosomes. Modified chromosome is obtained by cross over operation.
R K Jena[13]	Multiobjective nested particle swarm optimization	Maximum scheduling in less time.	Scheduling is done using nested PSO
M Kowsigan et al[14]	Fuzzy set and GA	Task scheduling with load balancing by reducing makespan	Scheduling is done by combining Fuzzy set with with GA which reduced the population
Ipsitakar [17]	GA	Task scheduling for minimizing energy consumption	Assigning of task is done by using max-min and GA algorithm
Huda Ibrahim[21]	Integer Linear Programming and adaptive genetic algorithm	Minimizes energy consumption ,optimally schedules the task	Dynamic scheduling with batch mode.

III. PROPOSED METHOD

In the proposed method an n number of tasks T1…Tn are given to the cloud service provider (CSP). Each task has different requirements in terms of resource, execution time, make span, deadline, frequency.

A. Cloud service provider -

It is responsible for providing the requests which are submitted by the users of the cloud. Each task will have different requirements. All these Tasks in the further step will be given to the scheduler.

B. Scheduler -

The scheduler allocates the different task to VM's. Each of the task is interpret with there requirements. A meta heuristic approach and ANN with back propagation is applied for scheduling the task efficiently. Meta heuristic algorithms are used for training ANN for efficient task scheduling.

Figure 3 the proposed model is given.

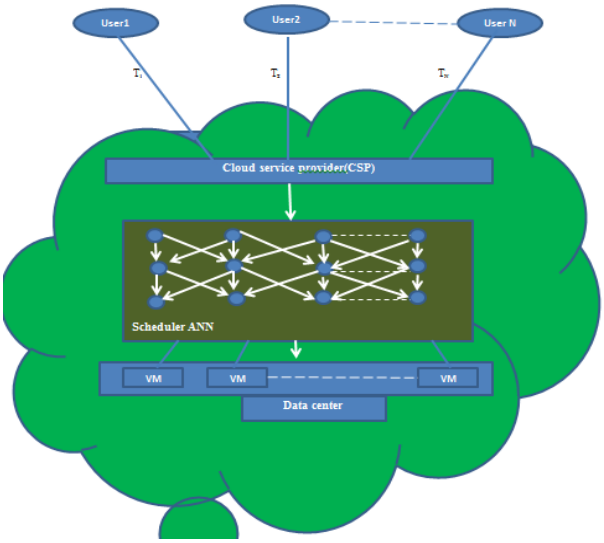

Figure 3: A Model of task scheduling

IV. FUTURE WORK

As efficient usage of energy in our concern in green cloud computing to reduce carbon footprints which is generated through datacentre. A comparative analysis of various algorithms for task scheduling using soft techniques is given and a model is proposed for efficient task scheduling using meta heuristic and ANN, assuming centralized data centers. The future work can be carried out for decentralized servers.

V. REFERENCES

[1]. Gan Guo-ning, HUANG Ting-Iei and GAO Shuai,"Genetic simulated Annealing Algorithm for Task Scheduling based on cloud computing environment', IEEE,2010

[2]. The NIST Definition of Cloud Computing Peter Mell Timothy Grance NIST Special Publication 800-145 T september 2011.

[3]. Linan Zhu, Qingshui Li and Lingna He, "Study on cloud computing a resource scheduling strategy based on the Ant Colony Optimization", International Journal of computer science",IJCSI,Volume.9, Issue 2, 2012.

[4]. Pradeep Kumar Amandeep Verma , "Independent Task Scheduling in cloud computing by Improved Genetic Algorithm", International Journal of Advanced Research in computer science and software engineering, Volume 2, Issue 5, 2012.

[5]. Jianhua Gu et al, "A new resource scheduling strategy based on Genetic Algorithm in Cloud Computing Environment", Journal of Computers, volume.7, No.1,2012
[6]. Medhat A Tawfeek and Arabi E. keshk, Fawzy A. Torkey,"Cloud task scheduling based on Ant colony optimization", IEEE, 2013.
[7]. Jing Liu et al, "Job Scheduling model for Cloud Computing based on multi objective Genetic Algorithm", International Journal of Computer Science Issues, Volume.10, Isuue 1, No.3, 2013.
[8]. https://www.datacenterknowledge.com/archives/2014/12/17/undertaking-challenge-reduce-data-center-carbon-footprint - 2014
[9]. Liji Jacob et al, "Resource scheduling in Cloud using Bacterial Foraging optimization algorithms", International Journal of Computer Applications, Volume .92, No.1,2014
[10]. Wanneng shu , Wei Wang and Yunji Wang, " A novel energy efficient resource allocation algorithm based on immune clonal optimization for Green Cloud Computing",EURASIP Journal on wireless communications and networking, 2014.
[11]. A.I. Awad et al," Enhanced particle swarm optimization for task scheduling in Cloud Computing environments, International conference on communication, Management and Information Technology, Elsevier, 2015
[12]. Mohammed Shojafar et al,"FUGE: A Joint meta heuristic approach to cloud job scheduling algorithm using FUZZY theory and a Genetic method", Springer, 2015
[13]. R K Jena,"Multi objective Task scheduling in Cloud Enivornent using nested Particle swarm optimization frame work, International Conference on Recent trends in Computing (ICRTC), Springer,2015
[14]. M Kowsigan and 2Dr. P. Balasubramanie, "Scheduling of Jobs in Cloud Environment using Soft Computing Techniques, "International Journal of Applied Engineering Research, Volume.10, No. 38, ISSN 0973-4562, 2015.
[15]. *Nima Jafari Navimipour and Farnaz Sharifi Milani,"* Task Scheduling in the Cloud Computing Based on the Cuckoo Search Algorithm", IJMO,Volume.5, ISSN: 2010-3697 ,2015
[16]. https://en.wikipedia.org/wiki/Power_usage_effectiveness - 2016
[17]. IPsitakar, R.N. Ramakant Parida and Himansu Das, "Energy aware scheduling using Genetic Algorithm in Cloud Datacenters", International Conference on Electrical, Electronics, and optimization techniques(ICEEOT), IEEE,2016
[18]. https://atfirstlight.net/key-components-of-a-data-centre/ 2017
[19]. https://www.energy.gov/eere/videos/energy-101-energy-efficient-data-centers 2017
[20]. kancharana Nithin Kumar and Richa vasuja ," A Novel scheme of computing : Green Cloud Computing", International Journal of Scientific Research in Computer Science Engineering and Information Technology(IJSRCSEIT),ISSN:2456-3307,Issue 3,2018
[21]. Huda Ibrahim , Raafat O.Aburukba and Khaled El-Fakih "A Integer linear Programming model and Adaptive Genetic Algorithm approach to minimize energy consumption of Cloud Computing Dataceners", Computers and Electrical Engineering, Elsevier, 2018.
[22]. Shatakshi Kaushal , Dweep Gogia and Bishwesh Kumar, "Recent Trends in Green cloud computing" ,Springer Nature Singapore, 2nd International Conference on communication, computing and networking, 2019