



Energy-Efficient Real-Time Tasks Scheduling in Cloud Data Centers

Debendra Maharana, Dr. Bibhudatta Sahoo, Dr. Srinivas Sethi

Department of Computer Science and Application, Rourkela Institute of Management Studies (RIMS), Rourkela, India

Department of Computer Science & Engineering, National Institute of Technology (NIT), Rourkela, India

Department of Computer Science & Engineering(IGIT), Saranga, Dhenkanal,India

deb.maharana@gmail.com , bdsahu@nitrkl.ac.in, igitsethi@gmail.com

Abstract—Reducing energy consumption in cloud computing systems has been a major concern among the researchers because it not only reduce the operational cost but also increase the system reliability, and efficient scheduling approach is a promising way to achieve this goal. But unfortunately, existing energy-aware scheduling approaches are inadequate for real-time tasks running in cloud environment because they assumes that cloud computing environment are deterministic and pre-computed schedule decisions are followed during the execution. The above issues are addressed in this paper by considering the number of energy-efficiency factors such as energy cost, CPU power efficiency, carbon emission rate, and workload, and near-optimal energy efficient scheduling policies are proposed for cloud data center for scheduling real-time, aperiodic, independent tasks that can reduce operational cost and provide Quality of Service (QoS).

Keywords—cloud computing; energy efficiency; real-time, scheduling;

I. INTRODUCTION

Cloud computing has emerged as a revolutionary paradigm for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources(e.g., networks, servers, storage, applications, and services) for customers based on a “pay-as-you-go” model[2]. Nowadays, an increasing number of enterprises and governments have deployed their applications including commercial business and scientific research in clouds. To meet such a huge demand of computing power, many industries are rapidly deploying distributed data centers containing large number of computing nodes in different locations around the world .

Data center industry has been growing rapidly in the last five years and reducing the energy consumption of

data centers has been a challenging issue because the data needed for computing are growing so quickly that increasingly larger servers and disks are needed to process them fast enough within the required time period.

According to Forrester research ,India will be the third largest internet user by 2015, with China and the US taking the first two spots respectively. The internet population is expected to grow by 10-20% annually for the next two years. With a growing internet population, the data to be handled is expected to grow exponentially. Infrastructure needs to support this kind of growth is also expected to grow leaving a strong forecast for the data center market.

With the rapid increase of the data centers, energy consumptions in data centers is also increasing at very rapid pace. Increasing energy cost imposes tremendous pressure on the developers to design energy efficient data centers. As per the Energy insights report-2015 by Govt of India, in business area and Industry part, the utilization of power has expanded at a much speedier pace contrasted with different segments amid 2005-06 to 2013-14 . Of the aggregate utilization of power in 2013-14, industry division represented the biggest share (43.83%), trailed by agribusiness (18.03%), and business parts (8.72%) .According to Industry Outlook 2014 , server farms devour up to 3% of all worldwide power creation, while producing more than 200 million metric huge amounts of CO₂ [16]. The most part of power consumption in data centers comes from computation processing, disk storage, network, and cooling systems. Beside the operating cost, huge consumption of energy in data centers result in low reliability of the system since the hosts failure rate doubles for every 10-degree increase in temperature[1].Therefore, reducing energy consumption has become a great challenge for operating cloud data centers .

In addition, the real-time task has been a primary fraction of computing applications in cloud data centers. For reducing energy consumption in real time system, two most widely used techniques are DVFS(Dynamic voltage and Frequency Scaling) and DPM(Dynamic Power Management).DVFS approaches reduce energy while improving performance by decreasing the voltage and the frequency of the CPU, but reducing the frequency increases the task execution time which cause a great challenge for scheduling real-time tasks efficiently in cloud system, because it need to derive maximum processor speed to meet the timing requirements/deadline of task while minimizing the energy consumption

Remainder of this paper is organized as follows. Section II presents the related work in the literature. Section III presents problem formulation includes the System , task and energy consumption model. The energy-efficient scheduling algorithm for real-time tasks is introduced in Section IV. Section V evaluated the performance of proposed algorithm by performing extensive experiments to comparing it with existing algorithms. The future work and the summary of conclusion is presented in Section VI.

II. RELATED WORK

In recent years, high energy consumption in cloud data centers has been a great challenges among the researchers, To address this issues, a large number of energy-aware scheduling approaches have been developed. Such as DVFS, DPM, machine virtualization etc. Among them, the DVFS technique has been widely used for reducing the power consumption of cloud data centers. For example,

In [3], Kyong Hoon Kim et al. have proposed several job scheduling methods for DVS-enabled cluster systems. The proposed policies consider the deadlines of tasks to regulate their execution speeds. In [11] Rodrigo N et al. proposed an approach for reducing the power required to execute urgent, CPU-intensive Bag-of-Tasks applications on cloud infrastructures. It exploits intelligent scheduling combined with the Dynamic Voltage and Frequency Scaling (DVFS) capability of modern CPU processors to keep the CPU operating at the minimum voltage level (and consequently minimum frequency and power consumption) that enables the application to complete before a user-defined deadline. . In [7], Anton Beloglazov et al. surveyed and classified various approaches to control the power consumption by a system from the OS level applying DVFS and other power saving techniques and algorithms.

However, the above studies show mild improvements on total energy consumptions of a data-center under their allocating policies.

In the global task scheduling aspect, several studies have focused on resource provisioning for the initial time. For this, bin packing heuristics have been proposed. for example,

In [4], Bo Li et al. uses an energy-aware First Fit bin packing-based algorithm. The migration frequency in the proposed method is sensitive to the mean arrival rate of tasks. In [5], Rajkumar Buyya et al. implemented an energy-aware VM manager that guides the reallocation of VMs according to current resource requirements and states of physical nodes . It uses modified Best fit decreasing (MBFD) bin packing - based algorithm which sort all VMs in decreasing order of current utilization and allocate each VM to a host that provides the least increase of power consumption due to this allocation.

In [6], Anton Beloglazov et al. proposed local scheduling policies typically have defined a number of thresholds for CPU utilization. In other words, they bound CPU utilizations of hosts not to be lowered than a minimum threshold nor exceed a maximum threshold In [9] Shailesh S. Deore et al. proposed a scheduling scheme for virtual machine called EESS(Energy-Efficient Scheduling Scheme) for reducing energy consumption by distributing maximum workload on minimum number of virtual machines. In [10] Xiaomin Zhu et al. used a rolling-horizon policy presented for enhancing the system's schedulability and presented an energy-aware scheduling algorithm EARH for real-time tasks. In [13] Seyedmehdi Hosseinimotlagh et al. proposed a cooperative two-tier task scheduling approach to benefit both cloud providers and their customers. It regulates the execution speeds of real-time tasks in a way that a host reaches the optimum level of utilization instead of migrating its tasks to other hosts. In [16] Tudor cioara et al. proposes a methodology for energy aware management of virtualized data centers (DC) based on dynamically adapting and scaling the computing capacity to the characteristics of the workload.

But unfortunately, to the best of our knowledge, few work considers the issue of dynamic energy-efficient real-time tasks scheduling in clouds environment.

III. PROBLEM FORMULATION

In this section, we will introduce the models, notions, and terminology used in this paper. Table 1 describes the different parameter used in the models.

Table 1. parameter description

Symbol	Description
H_a	active host set
h_i	i^{th} host
c_i	CPU performance measured in MIPS of i^{th} host
m_i	Memory size of i^{th} host
s_i	Storage capacity of host h_i
e_i	energy consumption in the host i
VM_i	Set of virtual machines for each host
v_{ik}	k^{th} virtual machine on i^{th} host
cv_{ik}	CPU capability of i^{th} virtual machine on k^{th} host
mv_{ik}	Memory size of i^{th} virtual machine on k^{th} host
sv_{ik}	Storage capacity of i^{th} virtual machine on k^{th} host
ev_{ik}	energy consumption of i^{th} virtual machine on k^{th} host
T	Task set
t_i	i^{th} task
ecn_{ijk}	Energy consumption by i^{th} task on j^{th} vm on k^{th} host

A. System Model

Figure 1 describe the cloud System model. The targeted system in this model is cloud data center consisting of n computing host $H = \{h_1, h_2, \dots, h_n\}$ where the active host set $H_a \subseteq H$.

For each host h_k , it is characterised as $h_k = \{c_k, m_k, s_k, e_k\}$, where c_k is the CPU performance of host h_k , m_k is the memory size, s_k is the storage capacity and e_k is the energy consumption of host h_k . For each host h_k , a set of virtual machine VM_k is denoted by $VM_k = \{v_{1k}, v_{2k}, \dots, v_{nk}\}$, where v_{ik} is the i^{th} virtual machine of host h_k .

B. Task Model

In this model, it is considered the independent, real-time, aperiodic task denoted by $T = \{t_1, t_2, \dots, t_m\}$. Where, $t_i \in T$ is the task submitted by the user and is modelled as $t_i = \{a_i, d_i, l_i, f_i\}$, where a_i, d_i, l_i, f_i represent the task arrival time, deadline, length of task and task finish time t_i respectively. The execution time of task t_i

on virtual machine v_{ij} is defined as $et_{ijk} = \frac{l_i}{cv_{jk}} \dots (1)$,

where, l_i is the length of task and cv_{jk} is the CPU capability of j^{th} virtual machine on k^{th} host.

C. Energy Consumption Model

In a cloud data center, since CPU consume the main part of energy, consumed by a host, so the energy consumption by CPU is considered in this paper. Further, the consumption of energy in cloud data center is classified into static energy consumption and dynamic energy consumption[10], and since static energy consumption is having less dominant than dynamic energy consumption, In this model, focus is made on the dynamic energy consumption.

Let the task t_i running on VM v_{jk} has consume the energy ecn_{ijk} , and let ecr_{jk} be the energy consumption rate of the VM v_{jk} . The energy consumption ecn_{ijk} can be

$$ecn_{ijk} = ecr_{jk} \cdot et_{ijk} \dots (2)$$

Since in a cloud data center the energy consumption is incurred when the VM is executing task and when VM is sitting idle, so the total energy E consumed by a host can be $E = ecn^{exec} + ecn^{idle} \dots (3)$.

Hence, the total energy consumed when VM executing all the tasks is:

$$ecn^{exec} = \sum_{k=1}^{H_a} \sum_{j=1}^{VM_k} \sum_{i=1}^T x_{ijk} \cdot ecn_{ijk} = \sum_{k=1}^{H_a} \sum_{j=1}^{VM_k} \sum_{i=1}^T x_{ijk} \cdot ecr_{jk} \cdot et_{ijk} \dots (4)$$

Where x_{ijk} represents tasks mapping to VM at different hosts in cloud data center, where x_{ijk} is "1" if task t_i is allocated to VM v_{jk} at host h_k and is "0", otherwise.

Similarly, when all the VMs in a host are sitting idle, by DVFS technology the host can be set to a lower energy consumption rate ecr_{jk}^{low} of virtual machine VM v_{jk} . In a host h_k , if the idle time is idt_k when all the VMs are idle, the energy consumption can be written as:

$$ecn^{idle} = \sum_{k=1}^{H_a} \sum_{j=1}^{VM_k} ecr_{jk}^{low} \cdot idt_k \dots (5)$$

IV. ENERGY-EFFICIENT SCHEDULING ALGORITHM

In the energy-efficient scheduling algorithm, the start time of task t_i on VM v_{jk} is represented as st_{ijk} and the ready time of VM v_{jk} is represented as rt_{jk} . If a new task t_n is allocated to VM, the new ready time on v_{jk} is,

$$rt_{jk} = st_{njk} + et_{njk} \dots (6)$$

Algorithm1: pseudo code of scheduling algorithm

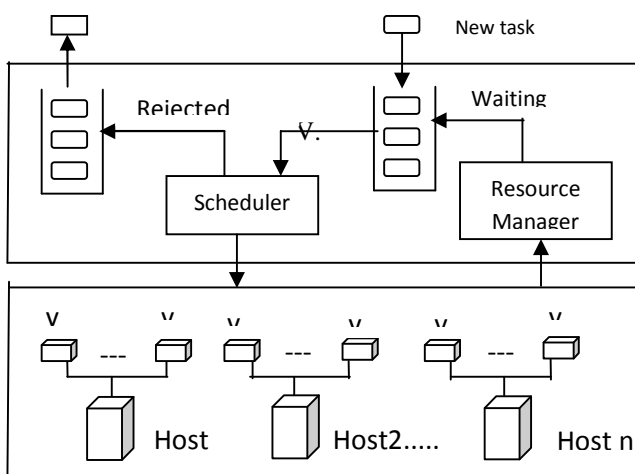
```

1:  flag ← FALSE; VMavail ← NULL;
2:  for each VM vjk in the host do
3:  Calculate the execution time etijl using Eq(1);
4:  if stijk + etijk ≤ di then
5:  flag ← TRUE;
6:  Calculate ecnijk using Eq.(3);
7:  end if
8:  end for
9:  If flag == TRUE then
10: Select vsk with minimum energy consumption
to execute ti; VMavail ← vsk; xijk ← 1; else
11: xijk ← 0;
12: end if
13: Update the scheduling decision of ti and
remove it from queue;

```

In the above scheduling algorithm, each task t_i is allocated to a VM, to meet their deadline, while minimizing the energy consumption. It first calculates the start time and execution time of task t_i on each VM before allocating it to VM. If it satisfies its deadline then the task is allocated to VM and then algorithm selects the VM that consumes the less energy to execute the task t_i , otherwise the task is rejected by the algorithm.

Figure 1. System Model



V. PERFORMANCE EVALUATION

The performance of the aforementioned algorithm is evaluated by using CloudSim toolkit[17] as a simulation platform. To evaluate its performance it is compared with three existing algorithms- EDF (Earliest Deadline First), FCFS (First Come First Serve) and EARH (Energy Aware Rolling-Horizon). The brief descriptions of these algorithms are as below.

FCFS: It assigns the new task as soon as it arrives to any available Cloud Resource.

EDF: Tasks are assigned to VM by their deadline in a non-decreasing order by guaranteeing the deadline of the task.

EARH: This algorithm employs the Rolling-Horizon Optimization scheme and Virtualization technique for scheduling the Real-Time task in Cloud.

The performance metrics, parameters and the detailed setting for conducting experiment are as given below.

Simulation parameters and detailed setting:

The CPU performance of each host with one CPU core is 500 MIPS, 1,000 MIPS, and 1,500 MIPS. Energy consumption rate of three different kinds of hosts are 250 W, 300W, or 400W. The task arrival rate is in Poisson distribution, and the parameter *intval* is used to determine the time interval between two consecutive tasks. The start-up time of a host is 60 s and the creation time of a VM is 10 s.

Performance Metrics:

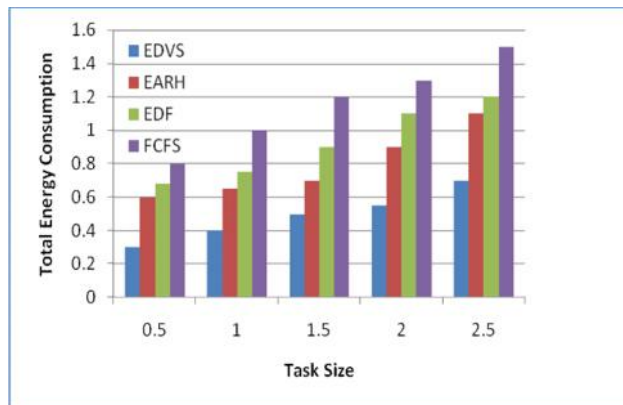
The various performance metrics used for evaluating system performance are,

ECPT (Energy consumption per task) = Total energy consumption / Task size.

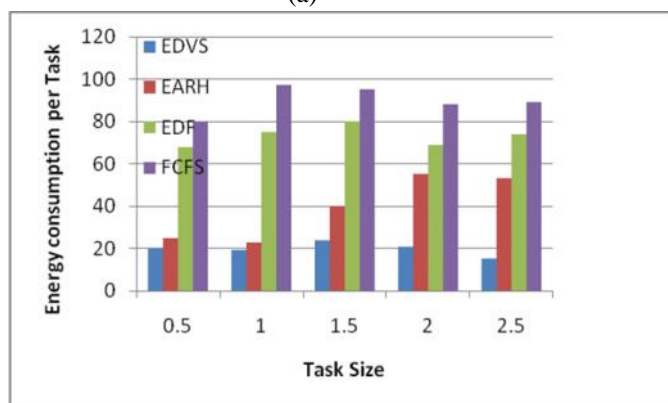
GR (Guarantee Ratio) = Number of tasks guaranteed to meet their deadline / Task size.

TEC (Total energy consumption) = Total energy consumed / Host

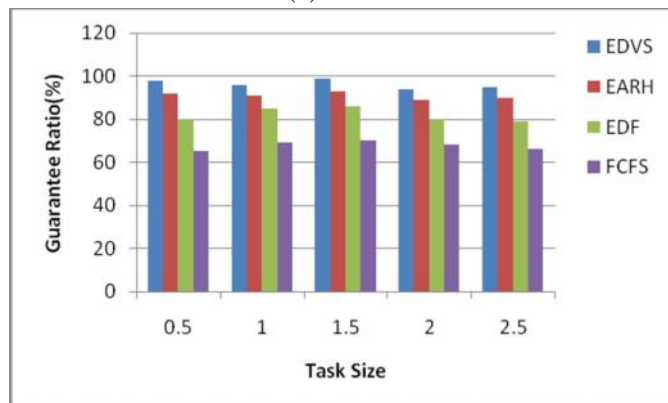
Different experimental results are presented below to observe the performance comparison of the algorithm in terms of the number of tasks executed.



(a)



(b)



(c)

From the Fig. a, it is observed that, the proposed algorithm EDVS has low energy consumption compared with others. When the task size is increased, the running VM can be consolidated to provide room for the new VM by avoiding energy consumption caused by the new active host. Further, the idle host can be shut down to reduce the energy consumption.

From Fig. b, it is observed that, energy consumption of EARH has slightly increased with increase of task size, whereas other three algorithms maintain stable energy consumption with increase of task size. This experimental result can be explained as, when task size increased, active VMs are consolidated to make room

for executing new tasks by reducing energy consumption of adding new tasks.

From Fig. c, it reveals that, all the algorithm maintain stable guarantee ratio regardless of increase in task size. This is because of the infinite resources in the cloud and new host and VM can be started with increase of task size. However, some task cannot finish execution successfully, because starting new host or adding new VM incur addition time cost, as a result some real-time task with tight deadline cannot complete their execution with the given time constraints.

VI. CONCLUSION

In this paper, energy-efficient scheduling problem for real-time independent aperiodic task in Cloud environment is investigated. The objectives of the scheduling of real-time independent tasks in Cloud computing system are to save energy consumption in the cloud data centers and to improve the system performance. To achieve the objectives, near-optimal energy efficient task scheduling policies are proposed for cloud data centers. The results of the experiment indicate the effectiveness of the proposed EDVS algorithm compared to related algorithm in cloud environment. In future research, it is planned to work on improving the execution time of real time task in the cloud environment.

REFERENCES

- [1] Kirk W. Cameron, Rong Ge, Xizhou Feng, "High-performance, power-aware distributed computing for scientific applications", vol. 38, Issue. 11, pp. 40-47, 2005.
- [2] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the Fifth Utility", Future Generation Computer Systems, vol. 25, no. 6, ELSEVIER, pp- 599-616, 2009.
- [3] Kyong Hoon Kim, Anton Beloglazov, Rajkumar Buyya, "Power-aware provisioning of Cloud Resources for Real-time Services" ACM 978-1-60558-847-6/09/11, pp. 1-6, 2009.
- [4] Bo Li, Jianxin Li, Jinpeng Huai, Tianyu Wo, Qin Li, Liang Zhong, "EnaCloud: An Energy-saving Application live placement approach for Cloud Computing Environments", IEEE International Conference on Cloud Computing, DOI 10.1109, pp. 17-24, 2009.
- [5] Rajkumar Buyya, Anton Beliglazov, Jemal Abawajy, "energy-efficient management of data center resources for cloud computing: A vision,

- architectural elements and open challenges”, Cloud Computing and Distributed Systems(CLOUDS) laboratory, the university of Melbourne,Australia, pp. 1-12, 2010.
- [6] Anton Beloglazov, Rajkumar Buyya, “Energy efficient resource management in virtualized Cloud data centers”, 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, DOI 10.1109/CCGRID.2010.46, pp. 826-831, 2010.
- [7] Anton Beloglazov, Rajkumar Buyya, Young Choon lee, Albert Zomaya, “A Taxonomy and survey of energy-efficient data centers and Cloud computing Systems”, ELSEVIER, advances in computers, vol. 82, pp. 47-111, 2011.
- [8] Toni Masteli, Ivona Brandi, “ Recent Trends in Energy Efficient Cloud Computing”, Journal of Latex Class Files, vol. 11, no. 4, pp. 1-6, 2012.
- [9] Shailesh S. Deore, Ashok Narayan Patil, “ Energy-Efficient Scheduling Scheme for Virtual Machines in Cloud Computing”, International Journal of Computer Application (0975-8887), vol. 56, no. 10, pp. 1-7, 2012.
- [10] Xiaomin Zhu, Laurence T. Yang, Huangke Chen, Ji Wang, Shu Yin, Xiaocheng Liu, “ Real-Time Tasks Oriented Energy-Aware Scheduling in Virtualized Cloud”, IEEE Transactions on Cloud Computing, vol 2, no. 2, pp. 168-180, 2014.
- [11] Rodrigo N. Calheiros , Rajkumar Buyya, “ Energy-Efficient Scheduling of Urgent Bag-of-Tasks applications in Clouds through DVFS”, IEEE 6th International Conference on Cloud Computing Technology and Science, DOI 10.1109/CloudCom.2014.20, pp. 342-349, 2014.
- [12] Seyedmehdi Hosseinimotlagh, Farshad Khunjush, Seyedmahyar Hosseinimotlagh, “A Cooperative Two-Tier Energy-Aware Scheduling for Real-Time Tasks in Computing Clouds”, IEEE 22nd Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, DOI 10.1109/PDP.2014.91, pp. 178-182, 2014.
- [13] Tarandeep Kaur, Inderveer Chana, “Energy Efficiency Techniques in Cloud Computing: A Survey and Taxonomy”, ACM Computing Surveys, Vol. 48, No. 2, Article 22, pp. 1-46, 2015.
- [14] Pragya, Manjeet Gupta, “ Analysis of Energy Efficient Scheduling Algorithms in Green Cloud Computing”, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), vol. 4 ,Issue 5, pp. 2413-2416, 2015.
- [15] Altaf Ur Rahman, Fiaz Gul Khan, Waqas Jadoon,” Energy Efficiency techniques in cloud computing”, International Journal of Computer Science and Information Security (IJCSIS), Vol. 14, No. 6, pp. 317-323, 2016.
- [16] Tudor Cioara , Ionut Anghel , Ioan Salomie, “ Methodology for energy aware adaptive management of virtualized data centers”, Springer Science, DOI 10.1007/s12053-016-9467-2, pp. 1-24, 2016.
- [17] R.N. Calheiros, R. Ranjan, A. Beloglazov, C.A.F.D. Rose, and R.Buyya, “CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms,” Software: Practice and Experience, vol. 41,no. 1, pp. 23-50, 2011.
- [18] J.O. Gutierrez and K.M. Sim, “A Family of Heuristics for Agent- Based Elastic Cloud Bag-of-Tasks Concurrent Scheduling,” Future Generation Computer Systems, vol. 29, no. 7,