Energy Efficient Task Scheduling in Cloud Computing

Prabhpreet Kaur (1), Monika Sachdeva (2)
(1) Research Scholar, Department of Computer Science Engineering, SBSSTC, Ferozepur
sandhuprabh36@yahoo.com
(2) Associate Professor, Department of Computer Science Engineering, SBSSTC, Ferozepur
monika.sal@rediffmail.com

ABSTRACT

Although cloud computing has been widely adopted by the industry, but the research on cloud computing is still at an infancy stage. Energy management is one of the challenging research issues. Cloud Infrastructure is the most important component in a cloud. It may consists tens of thousands of servers, network disks and devices, and typically serve millions of users globally. Such a large-scale data center will consume enormous amount of energy. This research presents a new approach of differentiating different types of workloads. The cloudlets submitted to the cloud provider are executed in an energy efficient DVFS approach.

Keywords


INTRODUCTION

Cloud computing is an area that is experiencing a rapid advancement both in academia and industry. This technology, which aims at offering distributed, virtualized and elastic resources as utilities to end users, has thepotential to support full realization of “computing as utility” in the near future [1]. Along with the advancements of the Cloud technology, new possibilities for Internet-based applications development are emerging. These new application models can be grouped in two ends: on one side, there are the cloud service providers that are willing to provide large-scale computing infrastructure at a price based primarily on usage patterns. It eliminates the initial high-cost for application developers of environment set up an application deployment. On the other side there are large-scale software systems providers, which develop applications such as social networking sites and e-commerce, which are gaining popularity on the Internet. These applications can benefit greatly of Cloud infrastructure services to minimize costs and improve service quality to end users. Previously, development of such applications required acquisition of servers with a fixed capacity able to handle the expected application peak demand, installation of the whole software infrastructure of the platform supporting the application, and configuration of the application itself. But the servers were underutilized most of the time because peak traffic occurs only at specific short time periods. With the advent of the Cloud, deployment and hosting became cheaper and easier with the use of pay-per-use, flexible elastic infrastructure services provided by Cloud providers. When these two ends are brought together, several factors that impact the net benefit of Cloud can be observed. Some of these factors include geographic distribution of user bases, capabilities of the Internet infrastructure within those geographic areas, dynamic nature of usage patterns of the user bases, and capabilities of Cloud services in terms of adaptation or dynamic reconfiguration, among others.

RELATED WORK

Yi-Ju Chiang et al. (2013) discussed that cloud computing is a new service model for sharing a pool of computing resources that can be rapidly accessed and released based on a converged infrastructure. In the past, an individual use or company can only use their own servers to manage application programs or store data. Thus it may cause the dilemma of complex management and burden in “own-and-use” patterns.

Lucio et al. (2014) presents a hybrid optimization model that allows a cloud service provider to establish virtual machine (VM) placement strategies for its data centers in such a way that energy efficiency and network quality of service are jointly optimized. Usually, VM placement is an activity not fully integrated with network operations.

Bharti Wadhwa et al. (2014) uses the carbon footprint rate of the datacenters in distributed cloud architecture and the concept of virtual machine allocation and migration for reducing the carbon emission and energy consumption in the federated cloud system. The proposed approach reduces the carbon dioxide emission and energy consumption of federated cloud datacenters as compared to the classical scheduling approach of round robin VM scheduling in federated cloud datacenters.

Fahimeh Farahnakian et al. (2015) investigated the effectiveness of VM and host resource utilization predictions in the VM consolidation task using real workload traces. The proposed approach provides substantial improvement over other heuristic algorithms in reducing energy consumption, number of VM migrations and number of SLA violations.

Sonika P Reddy et al. (2014) presented a system that handles real-time and non-real-time tasks in an energy efficient method without compromising much on neither reliability nor performance. Of the three processors, two processors i.e. the first and second, handle real-time tasks, using Earliest-Deadline-First (EDF) and Earliest-Deadline-Late (EDL) scheduling algorithms respectively.

Maurizio Giacobbe et al. (2015) presented a new strategy to reduce the carbon dioxide emissions in federated Cloud ecosystems. More specifically, we propose a solution that allows providers to determine the best green destination where virtual machines should be migrated in order to reduce the carbon dioxide emissions of the whole federated environment.
Samiran Roy et al. (2014) states that the computing is a computational framework that provides collection of virtualized resources as Service. Cloud computing is highly profitable cost effective services in the business world in the present day scenario. However, the energy consumption of Data Centers is the big problem emerging out of growing demand for cloud services.

Moona Yakhchi et al. (2015) presented that with rapid increasing demand of cloud computing technology, energy efficiency has become highly important in cloud computing infrastructures. Cloud computing concept offers low cost and high level of availability. However, it still has some challenging problems, such as resource management and power consumption.

Md Sabbir Hasan et al. (2015) discusses the proliferation of Cloud services have greatly impacted our society, how green are these services is yet to be answered. Although, demand escalation for green services has grown due to societal awareness, the approaches to provide green services and establish Green SLAs remain oblivious for cloud or infrastructure providers.

Chenxi Qiu et al. (2015) discusses that there is a need for cloud providers to optimize energy efficiency while maintain high service level performance to tenants, not only for their own benefit but also for social welfares (e.g., protecting environment). Both simulation and real-world

Yibin Li et al. (2015) states that Dynamic voltage scaling (DVS) has emerged as a critical technique to leverage power management by lowering the supply voltage and frequency of processors. In this paper, based on the DVS technique, we propose a novel Energy-aware Dynamic Task Scheduling (EDTS) algorithm to minimize the total energy consumption for smartphones, while satisfying stringent time constraints and the probability constraint for applications.

YunNi Xia et al. (2015) states that the increasing call for green cloud, reducing energy consumption has been an important requirement for cloud resource providers not only to reduce operating costs, but also to improve system reliability. Dynamic voltage scaling (DVS) has been a key technique in exploiting the hardware characteristics of cloud datacenters to save energy by lowering the supply voltage and operating frequency.

Mahesh B. Nagpure et al. (2015) describes that the cloud computing becomes well liked among cloud users by contribution of various resources. This is on demand as it gives dynamic flexible resource allocation, for reliable service in pay as use manner to cloud service users. Cloud computing is not application oriented and this is a service oriented. In cloud computing, dynamic flexibility in resource allocation propose by virtualization technology.

Y Edwin et al. (2015) proposes the concept which enhances the utilization of Green computing technique in servers and switches in order to reduce the overall power consumption in wireless storage area network. In this proposed system the linear power model and low power blade model were used to reduce energy consumption in central processing unit and calculation for the number of individual tasks carried out by the user and depending upon the user's requirement server is selected. The remaining servers are set to power management schemes such us voltage scaling, frequency scaling and dynamic shutdown techniques.

Research Motivation

Although cloud computing has been widely adopted by the industry, but the research on cloud computing is still at an infancy stage. There are many issues in Cloud computing such as Virtual Machine Migration, Data security, Energy Management, Server Consolidation etc. as discussed in previous section that have not been fully addressed. Energy management is one of the challenging research issues. Cloud Infrastructure is the most important component in a cloud. It may consists tens of thousands of servers, network disks and devices, and typically serve millions of users globally. Such a large-scale data center will consume enormous amount of energy. For example, according to research of Google datacenter used about 2.26 million MW hours of power to operate in 2010, resulting to carbon footprint of 1.46 million metric tons of carbon dioxide. In other words, a single data center can consume power which is equal to a power consumed by small town. In order to reduce power consumption, it is necessary to balance the load among the different nodes.

Green Computing is the practice of implementing procedures and policies that improve the efficiency of computing resources in such a way as to reduce the energy consumption and maintains environmental sustainability. Various existing scheduling techniques are there which manages load among the nodes but are not energy efficient for the Cloud computing platform. Aim of the thesis is to consolidate the load balancing in an efficient way so that the resource utilization can be maximized and the energy consumption of the data center could be minimized that can further result in reducing global warming and hence assist in achieving Green Computing.

METHODOLOGY

All the cloudlets created by the user are fetched by the Datacenter Broker. We have implemented the concept of Task Grouping. The tasks are grouped into 3 categories.

a. High computing task : Those tasks whose instruction length is greater than 1,00,000 are added to this group

b. Medium computing task : tasks have instruction length greater than 10,000 and less than or equal to 1,00,000.

c. Low computing task : Low computing tasks have instruction length less than 10,000.

We have taken 3 processors namely p1, p2, p3. They are arranged in the decreasing order of processing capabilities.
a. P1 will execute the high computing tasks.
b. P2 will execute the medium computing tasks
c. P3 will execute the low computing tasks.

If we are having a task with high instruction set and is running at DVS then it will consume more time, thereby consuming more power and emitting more CO2. If we are having a task with lesser instruction set and is running at Fmax then it will consume more power and there will be more CO2 emission. So we have scheduled the execution of tasks according to the following mechanism:

- High computing task will run at Fmax i.e. at 100%.
- Medium computing task will run at DVS i.e. at 60% frequency(approx).
- Low computing task will run at Threshold frequency i.e. at 40% frequency.

**ALGORITHM**

1. Create n number of cloudlets with different Instruction lengths.
2. Create 3 Virtual machines having different processors.
3. For all cloudlets C_i, i = 1 to n.
   - Compute Instruction length of each task. Divide the C_i into low, medium and high zone.
   - End;
4. For all Cloudlets in high zone
   - Assign C_i to P_1 processor that will be running at maximum frequency and First Come First Serve (FCFS) algorithm will be implemented.
   - End;
5. For all Cloudlets in medium zone
   - Assign C_i to P_2 processor that will be running at DVS and First Come First Serve (FCFS) algorithm will be implemented.
   - End;
6. For all Cloudlets in low zone
   - Assign C_i to P_3 processor that will be running at threshold frequency and First Come First Serve (FCFS) algorithm will be implemented.
   - End;
7. If any task on P_2 and P_3 fails, Schedule that task on P_1 i.e. the task will run at full frequency.
   - End;

---

**Figure 2. FlowChart of Proposed Work**
We have implemented First Come First Serve (FCFS) load balancing algorithm for High Computing tasks, Medium Computing tasks and Low Computing tasks on the same processor. If any task on P2 and P3 fails, schedule that task on P1 i.e. the task will run at maximum frequency. If no task has been assigned to the processor, then it will enter into the sleeping state to reduce the power consumption.

![System Design Diagram]

**EXPERIMENTAL SIMULATION**

We have used the CloudSim as a simulator for implementing the proposed methodology. Cloud service providers charge users depending upon the space or service provided. In R&D, it is not always possible to have the actual cloud infrastructure for performing experiments. For any research scholar, academician or scientist, it is not feasible to hire cloud services every time and then execute their algorithms or implementations. For the purpose of research, development and testing, open source libraries are available, which give the feel of cloud services. Nowadays, in the research market, cloud simulators are widely used by research scholars and practitioners, without the need to pay any amount to a cloud service provider. Different experiments are conducted several times on different number of cloudlets with different scenarios.

<table>
<thead>
<tr>
<th>Number of Cloudlets</th>
<th>Total Processing Time</th>
<th>Avg. Processing Time</th>
<th>Total Power Consumed</th>
<th>Total Processing Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>188600.00</td>
<td>3772.00</td>
<td>40545800.00</td>
<td>575418.60</td>
</tr>
<tr>
<td>75</td>
<td>277500.00</td>
<td>3700.00</td>
<td>59645000.00</td>
<td>846652.50</td>
</tr>
<tr>
<td>100</td>
<td>376300.00</td>
<td>3763.00</td>
<td>80951400.00</td>
<td>1148091.30</td>
</tr>
<tr>
<td>125</td>
<td>466100.00</td>
<td>3728.80</td>
<td>100190800.00</td>
<td>1422071.10</td>
</tr>
<tr>
<td>150</td>
<td>555000.00</td>
<td>3700.00</td>
<td>119290000.00</td>
<td>1693305.00</td>
</tr>
<tr>
<td>200</td>
<td>743600.00</td>
<td>3718.00</td>
<td>159835800.00</td>
<td>2268723.60</td>
</tr>
<tr>
<td>225</td>
<td>832500.00</td>
<td>3700.00</td>
<td>178935000.00</td>
<td>2539957.50</td>
</tr>
<tr>
<td>250</td>
<td>931300.00</td>
<td>3725.20</td>
<td>200241400.00</td>
<td>2841396.30</td>
</tr>
</tbody>
</table>

It is clear from the Figure 4, that the power consumption in the proposed work is very less as compared to the power consumed in the existing work. The values of different parameters like processing time, processing cost, power consumed by the different number of cloudlets have been calculated and the readings have been mentioned in Table 1. In the below bar chart, by comparing the energy efficiency, there is 23.9% reduction in power consumption. The user applications is based on the size of instruction (number of MIPS) of the applications.
The Figure 5 shows that as the number cloudlets increases the execution time also increases. As cloudlets i.e. number of instruction lines are increases the execution time also increases. This is the initial stage of the proposed algorithm in which it checks that if the demand of the users are increases the execution time also increases that means number of cloudlets are directly proportional to the execution time.

The processing cost occurred to the clients is directly proportional to the execution time of the cloudlets. As the number of cloudlets increases, their execution time will increase and thereby increasing the processing cost of the cloudlets as shown in figure 6.
CONCLUSION

This thesis gives the introduction of Cloud computing and background of various workload consolidation techniques to manage heterogeneous workloads. It also considers various existing load balancing scheduling algorithms. As the energy efficiency is one of major problem in cloud computing. So, in this work efficient energy consumption technique has been proposed. Many load balancing algorithms are existing today but no is one energy efficient as they balance the load among the nodes of virtual machines. The proposed technique can balance the load as well as it is energy efficient.

REFERENCES