International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8, Issue-11, September 2019

Energy-efficient Scheduling of Cloud Data Center Servers

M.Rudra Kumar, Z.Prathiba Rani

Abstract*: Data center is a cost-effective infrastructure to store large data volumes and host large-scale service applications. Providers of cloud computing services are deploying data centers worldwide quickly. With lots of servers and switches. These data centers consume substantial quantities of energy, which contributes to high operating costs. Optimizing server and network energy consumption in information centers can therefore decrease operating costs. In a data center, power utilization is chiefly because of servers, network devices, and cooling systems, an effective energy-saving strategy is to consolidate computing and communication into fewer servers and network devices and then power off as many unneeded servers and network devices as possible. A new method of reducing the energy utilization of computer systems and networks in data centers while meeting the requirements of the cloud tenants for quality of service (QoS) is proposed here in this paper.

Keywords*: Power Consumption, Server Power, Cloud Computing and Power Management.

I. INTRODUCTION

Data Center provides cost-effective storage of data facilities and large-scale services hosting. There are hundreds of thousands of data centers that are interconnected by switches, routers and high-speed connections. Big businesses, like Facebook, Yahoo!, Amazon and Google, regularly use information centers in storage, web-search and computing on a big scale.

Cloud computing, which provides computer resources as a service, is a Revolution technology that provides cost-effective and pay-per-use flexible IT use, namely networks, storage, servers, services and apps, without physically obtaining them [1]. This sort of computing offers companies with many benefits, little time for start fresh services, reduced the state of being preserved and costs of operating, greater use all the way through virtualization, and easier revival of disasters that make cloud computing an appealing option [2]. This technological advancement has allowed the creation of a fresh computing model in which resources (e.g., CPU and storage) are supplied as general utilities that consumers can lease and distribute via the on-demand fashion of the Internet. Cloud computing that enables customers to access on-demand services. According to the client request, it offers pool of shared data, software, databases and other devices resources. Cloud computing services relate to software, platform, infrastructure, information, identity and management of policies [3].

The cloud service delivery model in the cloud setting includes three primary services such as Infrastructure (IaaS), Platform(PaaS) and Software (SaaS) as a Service. Basic layer infrastructure services such as storage, database management and computing capacities are available on demand in IaaS. This platform used PaaS to design, create, construct and test apps. While SaaS is extremely scalable internet-based apps that are offered as end-user services where end-users can purchase and maintain overhead software or services supplied by SaaS [4].

In order to operate effectively the cloud data center, virtualization is a main technology. Resource for data centers is often underused as the average load of their ability is around 30 percent. In the virtualized information centers, energy consumption may be reduced by deciding which physical server should be placed on a virtual machine (VM). Virtual consolidation strategies attempt to host a certain number of virtual computers on as few physical computers as possible. According to the Project Report of Open Compute, a data center uses more than 93 percent of its energy consumption efficiently to use computer resources.

In the following figure: 1 Many data processing services that are handled by virtual machines are software as a service to be considered as a workload.

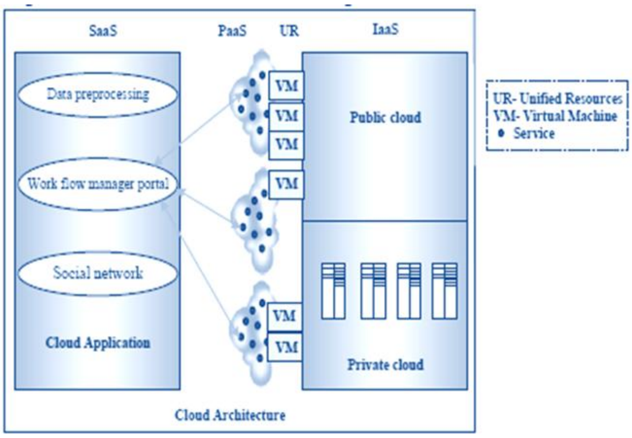

Fig 1: Cloud Architecture

Public cloud, personal cloud, community model and hybrid cloud are the four vital cloud organization models. Many computing service suppliers, including Yahoo, Microsoft, IBM and Google, are rapidly sending information centers to multiple places to deliver cloud computing services[5].

Revised Manuscript Received on September 05, 2019.
* Correspondence Author
M.Rudra Kumar, Professor & HOD, Department of CSE, Annamacharya Institute of Technology & Sciences, Rajampet, Andhrapradesh, India-516126.
Z.Prathiba Rani *, PG Student, Department of CSE, Annamacharya Institute of Technology & Sciences, Rajampet, Andhrapradesh, India-516126. Email: prathibarani521@gmail.com

The cloud computing has gone to the IT company with a particular end objective to boost high effectiveness and save energy through IT expansions. The global uptake of cloud computing has motivated drastic increases in the power consumption of the datacenter. Thousands of interconnected servers are comprised and operated through the datacenters to provide various cloud services.

The high energy consumption problem is becoming increasingly essential due to cloud computing technology and the development of big numbers of data centers fast developments. The information center's quality and effectiveness can be expressed in terms of the quantity of electrical power supplied [6].

Cloud computing involves a large number of data centers spread around the globe. Huge businesses like Yahoo, Google, Amazon and eBay are working on Cloud computing services with such big data centers. Data centers use huge amounts of energy, nevertheless. Greenpeace Report has stated that global demand for energy from the information center is estimated to be around 31 GW, or almost 180,000 households. The demand for electricity from data centers is growing rapidly. The power consumption in the United States in 2010 was between 1.7 percent and 2.2 percent of all the energy consumed by communication, power distribution, cooling, and servers. Environmentally producing CO_2 outcomes in global warming. Energy costs are increasing and natural resource depletion is increasing environmental concern.

II. RELATED WORK

For battery-operated system scaling, Dynamic Voltage and Frequency Scaling (DVFS) which is a voltage reduction method launched in the year 1990s[7], reduces energy usage utmost in big digital devices with the help of system's voltage and frequency adjustment for altering workloads. A regulated system equipped with DVFS can adjust a digital circuit's supply voltage at the functional boundary to the speed, temperature and parameters of the technology.

Sheshadri et al. [9] suggested heuristic approaches to minimize a SoC's overall testing time by scaling SOC's voltage and frequency below its power limit and individual cores ' maximum frequency limit.

As dynamic power is proportional to the supply voltage square, decreasing supply voltage can decrease active power consumption considerably. Multi-supply voltage (MSV)[8] is introduced to provide reduction of finer grain power and trade-off performance.MSV can create a chip with the smallest possible supply voltage to operate as slowly as possible. The voltage scaling method enables modules to use the lowest voltage amount on the critical routes, thus meeting the necessary timing limitations while enabling modules on non-critical routes to use reduced voltages, thus decreasing energy consumption. Compared to parallel architectures, this system tends to result in lower overhead region. As a result, two or more voltages of supply were used in the chip.

Dynamic power management relates to power management schemes introduced during the running of programs in electronic systems [9].Electronic systems can be seen as component collections that can be of a heterogeneous nature. Such components can be active at different times and consume various fractions of the power budget accordingly. On request, electronic devices are intended to produce maximum efficiency. However, peak performance is only required for certain time intervals.

Virtualization allows services to be transferred between servers and virtualization has several VMs that can use various multiplexed apps to share a server. Because the average amount of data center resources is about 30%, data centers can move virtual machines to consolidate workloads on a number of servers and then shut down underused servers, thereby saving excellent energy. The data center resources are underused. By choosing the VMs to be transferred on the heuristics associated with usage thresholds is the migration of VMs optimized In[10], they submitted a green energy-efficient planning algorithm that uses priority cloud computing schedule. For selecting VMs for executing employment, priority work schedule is used. The VMs are chosen based on the resource weight and SLA user requirements. Their technique can meet the minimum requirement for resources for a job and discourage excessive use of resources. The DVFS technology is applied in Cloud Computing to regulate server voltage and frequency. This method can decrease a server's energy consumption in idle mode or when the workload is light. Data Center Energy-Efficient Network conscious programming (DENS) aiming at equalizing data center energy consumption with efficiency, quality and traffic requirements. For consolidation allocation changes to the working loads, DENS is achieving this goal by implementing feedback channels between network switches to prevent any congestion or warm locations within the network that may definitely influence general efficiency. Overloaded switches can discourage congestion, which can lead to packet loss and maintain the elevated level of network usage of data centers. Recently, researchers have begun to explore energy-efficient scheduling with considerations of undeniable power consumption of leakage current for present and future manufacturing processes of circuits. Cores or voltage islands may be switched off whenever needed to save energy consumption.

III. EXPERIMENTAL SETUP

Therefore, the tool Cloud Reports20, takes for simulation the large-scale execution of distinct energy models on a real-world infrastructure. It utilizes Cloud-sim as a plug-in.

The simulation performed over one data center consists of 100 heterogeneous computer nodes (hosts designed to have one CPU core equal to 1000, 2000 or 3000 MIPS, 4 GB of RAM and 110 TB of storage capability. The simulation is conducted over a1-hour Infrastructure-as - a-Service (IaaS) Cloud platform. The IaaS simulated platform's clients are fully customizable. After simulation was over in the cloud reporting setting, in Matlab Neural Network Fitting Tool (NNFT) we introduced simulation outcomes to generate and train a network and assess its efficiency using MSE and Regression testing.

User and virtual machine (VM) parameters are set according to Table I and Table II.

A. Proposed Algorithm

Algorithm 1: Minimizing Total Server Power.

Input: R^t: set of requests in t, $R^t = \{r_1, \ldots r_{N^t}\}$;
S : set of servers, $S = \{s_1, \ldots s_M\}$.
Output: $\{x^t_{ij}\}$: the dispatching matrix for each requests;
$\{y^t_{j,0}, y^t_{j,k}, y^t_{j,k+}, y^t_{j,k-}\}$: servers' state in each time slot.
Initialization: all servers are in active state, so $y^0_{i,0} \leftarrow 1$.
for $t \leftarrow 1$ *to* T **do**
  Sort R^t by their end time in decreasing order;
  for $i \leftarrow 1$ *to* N^t **do**
    $S1 \leftarrow \varnothing$;
    for $j \leftarrow 1$ *to* M **do**
      if s_j *is active* && $s_j.cpu \geq r_i.cpu$ && $s_j.mem \geq r_i.mem$ && $s_j.disk \geq r_i.disk$ **then**
        $S1 \leftarrow S1 + \{s_j\}$;
      end
    end
    if $S1 \neq \varnothing$ **then**
      Select server s_j from $S1$ based on ***LongerNearest***;
      Allocate r_i onto server s_j, and $x^t_{ij} \leftarrow 1$;
      Update the available resources of s_j from time slot t to $(t + r_i.end - r_i.begin + 1)$;
    else
      $S2 \leftarrow \varnothing$;
      for $j \leftarrow 1$ *to* M **do**
        if s_j *is inactive* **then**
          $S2 \leftarrow S2 + \{s_j\}$;
        end
      end
      Select server s_j from $S2$ based on ***GreenSleep***;
      Backtrack and adjust the status of s_j;
      Allocate r_i onto server s_j, and $x^t_{ij} \leftarrow 1$;
      Update the available resources of s_j from time slot t to $(t + r_i.end - r_i.begin + 1)$;
    end
  end
  for $j \leftarrow 1$ *to* M **do**
    if s_j *is idle* **then**
      Sleep down the server s_j to sleep mode K;
    end
  end
end

Table-I: User Parameters

Users Parameters	
No.of Customers	5
Cloudlets sent per min	250
Avg.Length of Cloudlets	50000
Avg.Cloudlet's file size	500

Table-II: Virtual Machine Parameters

Virtual Machine Parameters	
No.of VMs	5
Avg.Image Size	1000
Avg. RAM	512MB
Avg. Bandwidth	100000Hz

IV. EXPERIMENTAL RESULTS

Figure 2 shows the proposed model's overall use of the CPU. The highest use of resources for proposed model is 1.0 MIPS from simulation. The average use of the CPU is 0.5 MIPS.

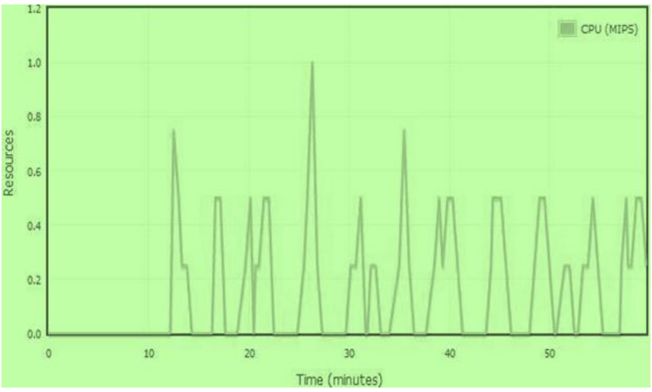

Fig 2: Overall usage of CPU

Figure 3 shows the total proposed model power consumption. During the 35th, 40th and 55th minutes, the highest value for power consumption for the Square root model is 1.75 KW from simulation. The average power consumption is approximately 0.8 KW.

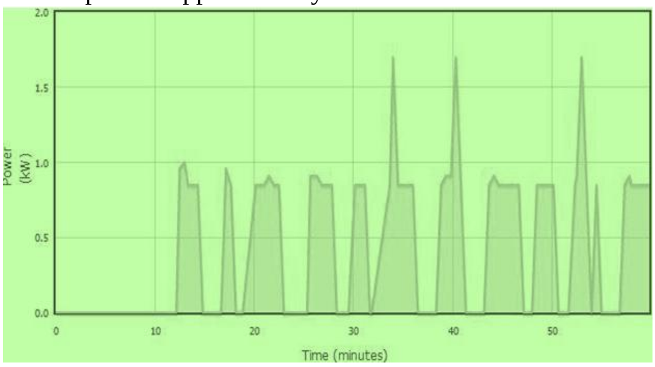

Fig 3: Proposed Model Power Consumption

V. CONCLUSION

The resources of the cloud computing datacenter consume a great deal of energy. To reduce power consumption and heat generation for green computing VM consolidation in data center plays a crucial role in optimizing resource use where overloaded hosts migrate the VMs to proper host servers as well as the under loaded host to be placed in idle mode in order to reduce power consumption. We assessed the efficiency of various energy models over IaaS Cloud infrastructure in this research paper. In addition, we discovered in each case the connection between CPU use and power consumption.

REFERENCES

1. R.Sood, R. Sharma "Cloud Security Threats and Issues-A Review", International Journal of Computer Sciences and Engineering,Vol.5, Issue-4, pp.115-119, April 2017.
2. Karthik Kumar and Yung-Hsiang Lu, "Cloud Computing for Mobile Users: Can Offloading Computation Save Energy", Published by the IEEE Computer Society, 2010.
3. Fatma A. Omara, Sherif M. Khattab and RadhyaSahal, "Optimum Resource Allocation of Database in Cloud Computing", Egyptian Informatics Journal, vol. 15, pp.1–12, 2014.
4. Saurabh Kumar Garg, Steve Versteeg and Rajkumar Buyya, "A framework for ranking of cloud computing services", Future Generation Computer Systems, vol. 29, pp. 1012–1023, 2013.
5. M. Abdullahi, M. Ngadi and S. Abdulhamid, "Symbiotic Organism Search optimization-based task scheduling in cloud computing environment", Future Generation Computer Systems, vol. 56, pp. 640-650, 2016.
6. Thai HaNguyen, Matthew Forshaw, Nigel Thomas " Operating Policies for Energy Efficient Dynamic Server Allocation" 318 (2015) 159–177.
7. Vijay Sheshadri, Vishwani D. Agrawal, Prathima Agrawal, "Power-Aware Optimization of SoC Test Schedules Using Voltage and Frequency Scaling" Journal of Electronic Testing , Issue 2/2017.
8. Raje, S. and Sarrafzadeh, M., "Variable voltage scheduling," Proceedings of the 1995 international symposium on Low power design, ACM, 1995, pp. 9–14.
9. Suh, J., Kang, D.-I., and Crago, S. P., "Dynamic power management of multiprocessor systems," Parallel and Distributed Processing Symposium, International, Vol. 2, IEEE Computer Society, 2002, pp. 0097b–0097b.
10. Benini, L., Bogliolo, A., and De Micheli, G., "A survey of design techniques for system-level dynamic power management," Very Large-Scale Integration (VLSI) Systems, IEEE Transactions on, Vol. 8, No. 3, 2000, pp. 299–316.

Retrieval Number: K17590981119/2019©BEIESP
*DOI: 10.35940/ijitee.*K1759*.0981119*

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication