RESEARCHERID
THOMSON REUTERS

[Jason, 6(1): January 2019] ISSN 2348 – 8034
DOI- 10.5281/zenodo.2532816 Impact Factor- 5.070

GLOBAL JOURNAL OF ENGINEERING SCIENCE AND RESEARCHES

COMPARATIVE STUDYOF SCHEDULING OF ENERGY EFFICIENCY IN CLOUD DATA CENTERS

Sebagenzi Jason[*1] & Suchithra. R[2]
[*1]Research scholar, Jain University, Bangalore560043, India
[2]Head of Department of MSc IT, Jain University, Bangalore-560043, India

ABSTRACT

A long time ago, cloud data centers accounted problems of offline and online process. This paper, is oriented on elaboration, utilization and comparison of offline and online scheduling which is the crucial point in all the systems. The conception and comparative study has been developed, and algorithms of scheduling are demonstrated in detail in the present paper. At the end, the analysis and the performance evaluation of those algorithms by simulation are also shown in order to be used.

Keywords- *scheduling Virtual machines, offline algorithm, online algorithm, ETF algorithm, Virtual machine, competitive ratio*

I. INTRODUCTION

At the present time, the industry related to the cloud computing seems to have the great importance after other resources we use in our daily activities.

Cloud computing is an architecture of the information technology services and business in the computing area, which distribute jobs to various data centers which are composed by a large numbers of virtual and also physical servers, and for that, programs applied to the system can be storage in dedicated spaces.

Cloud computing is till now in the stage of improvement, which incorporate also new technology of Cloud data centers philosophy which becomes very large according to the technology incorporate in it.

As an extra factor, energy consumption is increasing gradually due the different components incorporate in it: the cooling system, virtual physical machines etc.

The different studies demonstrated that the current consumption of the energy in Gross Domestic Product in China is very high than Japan, which is highest also than the United States and highest than Germany and France.
In 2012, In Chine country, cost of consumption of electricity, was around 14.00 billion by year, which is equal to the amount of power produced by China in Energy Engineering Corporation group called Harbin Electric power station.

A data center which has around 650 servers will cost around 2.0 million of bills during one year.
The use of the electricity by one server equipment which include the air-conditioning for cooling is similar to that used by the physical machine itself
.
In case 2 W of IT electricity is used, in the cooling system, more than 2 W of electricity is needed as well; however, in order to save 2 W of IT electricity, it will be necessary to reduce the amount of electricity used in the cloud computing.

For reducing the costs of the electricity consumption, cloud computing are forced to find means to use efficiently resources of data center and for that, it is possible to reduce electricity consumption.

By seeing the studies of statistics done by researchers, one data center, can use around 15% of its power given and this at any time.For that, 85% of its resources are wasted or idle.
Additional to that, only about 5% of the consumption of the electricity, is used for the processing of the data.

Today, by continuously improving Cloud computing, they must be possibilities of large-scale service scheduling.

Algorithms which are very well designed can: use and manage virtual resources and physical server in Cloud data centers dynamically, and they can provide resilient and flexible good jobs as well as services that are necessary for building good architecture of data centers, able to manage business, facilitate enterprises to finish realization of targets and reaching the level of using the electricity efficiently and also providing high performance services.

II. BACKGROUND

In order to determine exactly the electricity used in a computing system, the companies often the technique of indexing.

This techniques gives the percentage of electricity consumed by different devices which compose all the system.
Electricity consumed in data center comprises the consumption of all devices which are in the data center and the electricity which is used for different control.

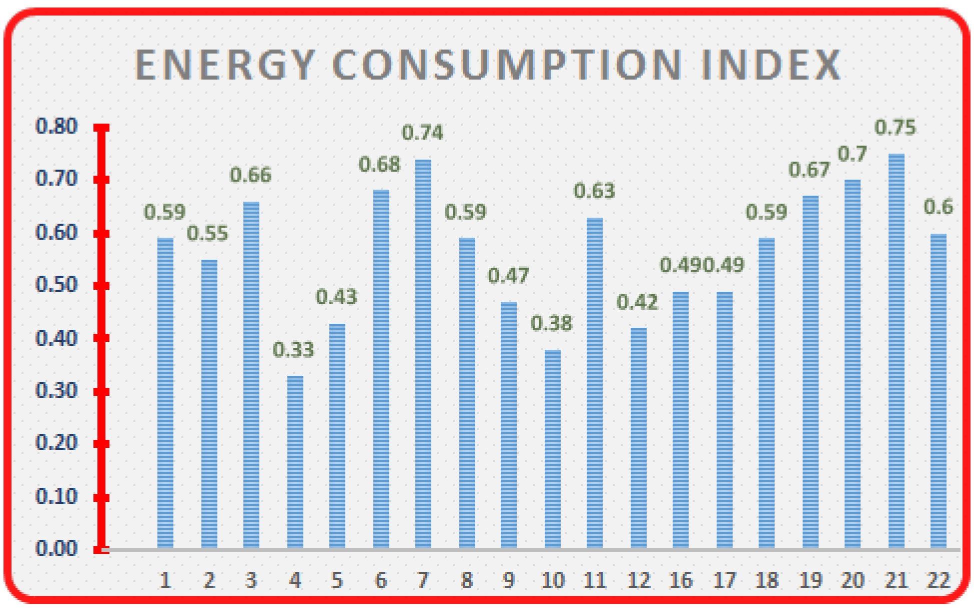


Figure1: The comparison of energy consumption index

The highest index shows good consumption of the power. The best target to reach in term of value of the index should be included between the value of 0.7 and the value of 0.8.

Figure 1 is showing clearly, 22different data center and is indicating the electricity index of every data center, and clearly, shows that the consumption of electricity in data centers has to be improved in order to reach good results.
Different algorithms must dynamically be put in place for resolving those problems.
Different researchers has done studies in order to reach that need.

III. RELATED RESEARCH

Beloglazov et al. [2] and Megha et al. [14]has made an introduction of cloud computing system status.
Bohrer[4] and Mathew et al. [13] discuss green computing;
Bollen et al. [5] and Rabiatul et al. [17]shows the background of research and challenges found in Cloud computing and energy problems.
Bonabeau[6] gives a detailed on the mechanism of green power-saving in the area of networks.
Coffman[7] make an analysis of the current status and progress of the electricity use in cloud.
Elnozahy and his colleagues [8] and Mofijul et al. [15]give too many scheduling algorithms for Cloud platforms.
Elnozahy and his colleagues [8] give too many several scheduling algorithms in that area.
Flammini and his colleagues [9] and Lefurgy and al. [12] discuss the cluster technology

Khargharia et al. [10] implement the platform in order to help users to manage the resources of the cloud data centers.

IV. STUDY OF OFFLINE ALGORITHMS

A. Energy Model for Data Centers

1. Evaluation of Data center consumption.
Different components which constitute the cloud data center, is a big quantity of physical servers as well as virtual machines.

Different type of servers are there, some having a single-core and others having multi-core processor.
Equipment situated in a single chassis is able to use only one single power supply device.
The system of cooling the cold and the heat are there and are situated between rows [11].

Air-conditioning is installed to give the cold air to the equipment, while the hot air is crossing through under-floor channels towards the ground.
The total consumption of the electricity is given by calculation processing, system of storage and the system for managing the temperature.

The formula gives the total consumption of the electricity in the system:

$$\boldsymbol{P_{total} = P_{pm} + P_{AC} + P_{additional}} \quad (8.1)$$

Where P_{total}designs the sum of the consumption of the electricity,$\boldsymbol{P_{pm}}$shows sumof consumption of physical servers, $\boldsymbol{P_{AC}}$shows how much the cooling system is consumed, and $\boldsymbol{P_{addditional}}$shows other consumptions in the whole system.
Consumption of the cooling system can be demonstrate using the equation [4]:

$$\boldsymbol{P_{AC} = \frac{P_{pm}}{C_o P(T_{sup})}} \quad \textbf{(8.2)}$$

Where $\boldsymbol{T_{sup}}$ is the quantity of temperature entering in air-conditioning and $\boldsymbol{C_o P(T_{sup})}$ is the factor which is showing the quality of the system.
The factor shows the consumption of the system of refrigeration by doing the comparison of the consumption of air-conditioner.

2. Electricity consumption model of machines
Utilization of a machine is related to the functionality and also the type of the micro processor in corporate in it.
By assigning $\boldsymbol{P_{min}}$, the utilization of the electricity when the machine is not working; and $\boldsymbol{P_{max}}$the electricity used when the server is underutilization, the following formula can be used:

ISSN 2348 – 8034
Impact Factor- 5.070

$$P = P_{min} + (P_{max} - P_{min})U \qquad (8.3)$$

Where P indicates the electricity used by the physical machine in cloud data center and U indicates the CPU use of the physical servers.

In some situation, use of the machine, time can influence the changes because of the variation of the workload. For that, microprocessor used is related to the time and is represented as$U_i(t)$.

For that, E_i which indicate the consumption of electricity by a physical machine, can be represented by:

$$E_i = \int_{t0}^{t1} P(U_i(t)dt \qquad (8.4)$$

If $U_i(t)$ is a fixed number over, then$U_i(t) = U_i$.

The consumption of the physical machines can be written as follows:

$$E_i = E_{i_{on}} + \sum_{j=1}^{k} E_{vm_j}$$

$$= P_{min}T_i + (P_{max} - P_{min)} \sum_{j=1}^{k} u_j t_j \qquad (8.5)$$

WhereT_i is the total of electricity when the machine is on,$u_j$indicates the utilization of virtual machines placed on physical machine. The value of time that VM_j works on physical machines is indicated by t_j.

The consumption can be calculated as following:

$$E_{cdc} = \sum_{i=1}^{n} E_i \qquad (8.6)$$

It is the sum of electricity utilized by physical machines of the system.

B. Algorithms in Offline State

By knowing information's related to the allocation of requested virtual machine and the need of electricity in the entire system, following algorithms are used:

FF algorithm

The target of FF (First Fit) algorithm, consist of sorting all VMs in no decreasing order.

The total load is a technique used to know the minimum physical machines needed in the system.

Then, by placing them in that way, the energy consumed by of all physical machines can be known easily.

This is the aim of algorithm 1

Input: VM requests indicated by their (required VM type ID_s , start time, ending-time, request capacity), the number of the request I is denoted as R_i.

Output: ID_s of PM_s for all VM_s, the number of the needed PM_s, the total energy consumption.

1. Sort the virtual machine in ascending order of their start-time;
2. For i= from 1 to n do
   d=0;
   a. If they are not overlapped or overlapped but still can share resources of an PM do
      Allocate i to the PM d;
      Else;
      Start a new PM; d=d+1; allocate i to PM d;
   b. End;
3. End for;

Algorithm 1:FF algorithm

1. *ETF algorithm*

The ETF is an algorithm which is focused mainly at the ending-time first (ETF). In that process, the virtual machines are sorted in no decreasing order. Based on the total load technique, it is very simple to find number of physical machines needed.

Then VMs are placed according tithe physical machine load

Input: VM requests indicated by their (required VM type ID_s , start time, ending-time, request capacity), the number of the request I is denoted as R_i.

Output: ID_s of PM_s for all VM_s, the number of the needed PM_s, the total energy consumption.

Initialization : allocating an ID to each PM.

1. Sort the virtual machine in ascending order of their ending-time;
2. For i= from 1 to n do
   d=0;
   If they are not overlapped or overlapped but still can share resources of an PM do
   Allocate i to the PM d;
   Else;
   Start a new PM; d=d+1; allocate i to PM d;
   End;
3. End for;

Algorithm 2: ETF algorithm

Consumption of energy of Offline algorithms is very high as the figure 2 is showing.

More the maximum duration is increasing and more the consumption of the energy is increasing also

[Jason, 6(1): January 2019] ISSN 2348 – 8034
DOI- 10.5281/zenodo.2532816 Impact Factor- 5.070

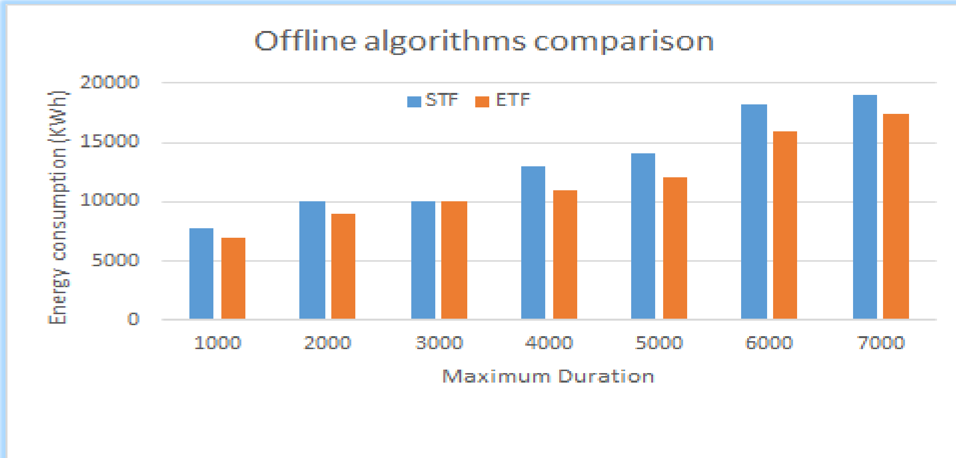


Figure 2: Offline Algorithm Comparison.

C. Algorithms in Online State

1. MFFD (Modified First-Fit Decreasing) algorithm

When Virtual Machine is requested, it must be placed automatically. The scheduler equipment should give information immediately to the user to which physical machine a request has been be assigned.

Input: VM requests indicated by their (required VM type ID_s , start time, ending-time, request capacity), the number of the request I is denoted as R_i.

Output: ID_s of PM_s for all VM_s, the number of the needed PM_s, the total energy consumption.

Initialization: allocating an ID to each PM.

```
1.  For i= from1 to n do
2.  d=0;
3.  when there comes the request of the VM do;
4.  If they are not overlapped or overlapped but still can share resources of an PM do
            Allocate i to the PM d;
    Else;
            Start a new PM; d=d+1; allocate i to PM d;
    End;
5.  End for;
```

Algorithm 3: MFFDE Algorithm

2. Bipartition-first-fit algorithm

This type of algorithm is based on schedules the job on principle of first-come-first-service, which consist of partitioning the time dynamically into two sub windows (bipartitioning).

The action of placing the request to the first machine that can host it is then taking place (first-fit).

ISSN 2348 – 8034
Impact Factor- 5.070

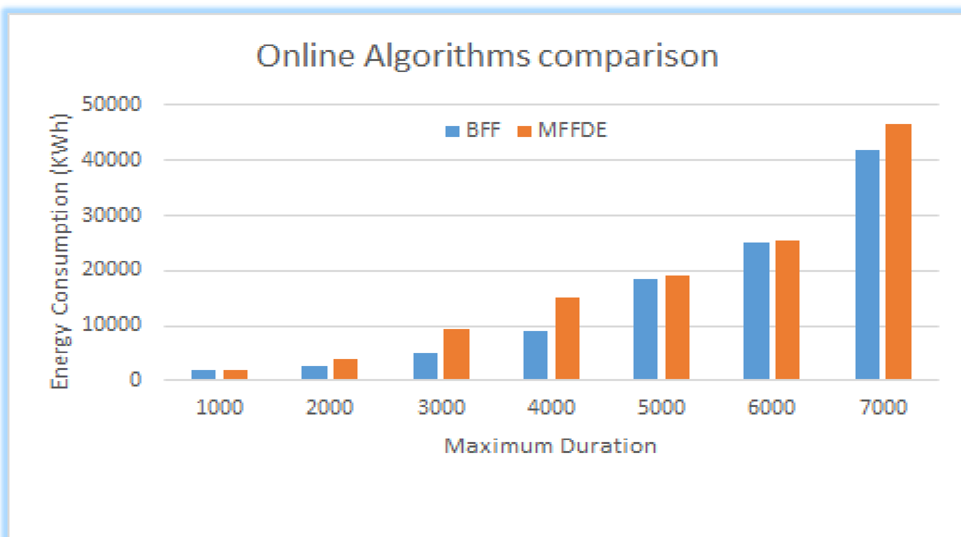
Input: g, the max capacity of a machine, and job I_i one by one.

Output: The scheduled jobs and total busy time of all machines and total number of machines used.

1. Allocates the first job to machine m_1;
2. For all job comes and time period in T do
   - a. If there is only one job,
     Allocate it to the first machine;
     Else;
     Computes the longest and the second longest interval of all current request in the system, set:
     $$k = \frac{span(\,the\ longest\ interval)}{span(the\ second\ longest\ interval)};$$
   - b. If $k > 1$,
     dynamically partitions the time plane into two sub windows LEFT and RIGHT using the median of all end-time of current requests as the partitioning point. The first job is counted as in LEFT window. Any job interval with end-time as left time of the partitioning point belongs to LEFT window, others belong to RIGHT;
     Else;
     Considers all jobs in LEFT window;

   End;
3. Counts workload and busy time of all machines;
4. Returns the set of machines used and total busy time.

Algorithm 4: bipartition-first-fit Algorithm



V. CONCLUSION

In this paper, the different type of algorithms related to offline and online situation has been shown and explained in detail.
However, as shown in References. [10, 16], different formulas has been shown for calculation of electricity consumed in the cloud data centers.

Graphics has been shown in order to compare different algorithms which can be used in the system.
These algorithms are compared between them in order to determine the efficiency of them.

VI. FUTURE WORK

There are still some challenges which can be seen in the future by researchers:
Researchers should focused on limitations of online and offline algorithms in order to get good solutions.
In this paper, we focus some algorithms. We believe that the results found can be extended for any number of jobs and still the results must be valid.

VII. ACKNOWLEDGMENT

My acknowledgment is addressed to anyone who contributes and gave the input and consistent explanations on the improvements of this paper.

REFERENCES

1. *Baker, Thar, Asim, Muhammad, Tawfik, Hissam (2017). An energy-aware service composition algorithm for multiple cloud-based to applications. J.Netw. Comp. App. 89 (1), 96-108.*
2. *Beloglizov, A., Abawajy, J., Buyya, R., 2012. Energy-aware resource allocation Heuristics for efficient management of data centers for cloud computing. Future Generat.Comput. System. 28 (5), 755 - 768.*
3. *Beloglizov, A., Buyya, R., Lee, Y.C, Zomaya, A.Y., 2011. In:Zelkonwitz, M. (ed), A Taxonomy and survey of Energy-efficiency Data Centers and Cloud Computing Systems, Advances in computers, vol.82. Elsevier, Amsterdam, The Netherlands, PP 47 - 111.*
4. *Bohrer, P., Elnozahy, E., Keller, T., Kistler, M., Lefurgy, C., McDowell, C., & Rajamony, R. (2014). The case for power management in Web servers. Norwell, MA, USA: Kluwer Academic Publishers.*
5. *Bollen, J., & Heylighen, F. (2015). Algorithms for the Self-organization of Distributed, Multiuser Networks. Austrian: Cybernetics and Systems.*
6. *Bonabeau, E., Dorigo, M., & Theraulaz, G. (2015). swarm Intelligence. USA: Oxford University Press.*
7. *Coffman Jr., E.G., Garey, M.R., Johnson, D.S., 1987. Bin-Packing with divisible item sizes. J. Complex 3 (1987), 406 - 428.*
8. *Elnozahy, E., Kistler, M., & Rajamony, R. (2015). Energy- Efficient server clusters. Power-Aware Computer Systems.*
9. *Flammini, M., Monaco, G., Moscardelli, L., Shachnai, H., Shalom, M., Tamir, T., Zaks, S. 2010. Minimizing total power-on time in parallel scheduling with application to optical networks. Theor. Comput. Sci. 411 (40 - 42), 3553-3562.*
10. *Khargharia, B., Hariri, S., & Yousif, M. (2015). Autonomic power and performance management for computing systems. USA: Cluster Computing.*
11. *Kim, K., Beloglazov, A., Buyya, R., 2011. Power- aware provisioning of virtual machines for real-time Cloud services. Concurrency Comput. Pract. Exp. 23 (13), 1491-1505.*
12. *Lefurgy, C., Rajamani, K., Rawson, F., Felter, W., Kistler, M., & Keller, T. (2015). Energy Management for commercial servers. USA: Computer 36 (12).*
13. *Mathew, V., Sitaraman, R.K., Shenoy, P., 2012. Energy-aware load balancing in content delivery networks. In: Proceedings of INFOCOM 2012, 25-30 March, PP. 954-*
14. *Megha, Desai R.; Hiren, Patel B.;. (2015). Efficient Virtual Machine Migration in Cloud Computing. Fifth International Conference on Communication Systems and Networking Technologies.*
15. *Mofijul, I. M. (2015). A Genetic Algorithm for Virtual Machine Migration in Heterogeneous Mobile Cloud Computing. Bangladesh.*
16. *Pankajdeep, K., & Rani, A. (2015). Virtual Machine Migration in Cloud Computing. Internationale Journal of Grid Distribution Computing Vol.8.*
17. *Rabiatul, A., Ruhani, R. A., Norliza, Z., & Mustaffa, S. (2015). Virtual Machine Migration Implementation in Load Balancing for cloud computing. Mara (UiTM).*