FP- International Journal of Computer Science Research (IJCSR)
Volume 1, Issue 1, Pages 17-20, 2014

MEUD: Minimization of energy utilization in the Cloud Data Center

Nataraj J
School of Information Technology and Engineering,
VIT University, Vellore-14, India
natarajk49@gmail.com

Muralikrishnan P
School of Information Technology and Engineering,
VIT University, Vellore-14, India
muralikrishnan1190@gmail.com

Praveen Kumar J
School of Information Technology and Engineering,
VIT University, Vellore-14, India
praveenkumar.j2013@vit.ac.in

Priya V
Assistant Professor (SITE)
VIT University, Vellore-14, India
vpriyacse@vit.ac.in

ABSTRACT

In computing environment of cloud plays vital role in the data center to share information over the communication medium. The cloud data centers are used to store and retrieve the large amount of data across the network with sufficient amount of energy tilized in the data processing system. In the prior work to be conducted using dynamic capacity provisioning method to save power and fine-tuning of the number of dynamic state machine for the requisite cloud utility demands. It might not be fully handle the heterogeneous cluster data center in the production environment. The data center has different capacities, capabilities and energy consumption is needed. Based on the analysis, propose the Heterogeneous Energy Utilization method to minimize energy consumption in the data center. Here we use Expectation Maximization algorithm to sort the nodes in each cluster in the data center. Clusters are enabled dynamically when the resource request to the data center, show simulation of saving energy and data center utilization effect.

Keywords

Data center, cloud backbone services, clustered nodes and power consumption factor.

1. Introduction

Cloud computing is the imperative development of Information technology to gain large-scale of data distribution system. The cloud service maintains huge number of servers and their interlinked data path. This service can be accessed using data centers. There are N numbers of servers using in the data distributing system. Data centers can maintain heterogeneous servers over the network. Many service providers are investing their shares and extend the scalability of organization infrastructure.

While accessing the data through data center[1], the energy dissipation is more as well as power consumption of each data center is increased, get more dissipation of non-renewable energy affect the organization infrastructure. The cost-effective of data centers increased, need to perform additional support and reduce the power dissipation to tolerate demanded outage of cloud resource. In each data center can maintain certain number of clustered nodes to access the data. If user requests some data from data center, active clustered node response their request immediately. The remaining number of clustered nodes in the data center acts as disabled condition. We propose the Heterogeneous Energy Utilization method can reduce the power dissipation and utilization factor of data center.

The section I explain the concept of the paper, section II is the background survey, section III is the power consumption, section IV is the proposed system, section V is the simulation result, section VI is the conclusion and section VII is the reference.

2. Background

The effective handling of data centers using heterogeneous workload scheduling method to minimize energy-consumption while convention of service level objective (SLO). Another metric of this method to manage the task scheduling in data centers. To adjust the location of each clustered node while it's in active state. This could avoid the workload scheduling delay.

Burak Kantarci and Hussein T. Mouftah [2] explain the usage of cloud data centers as pay per use. Data centers could active state when it ready to provide service. Minimum outcome of provisioning an efficient cloud framework reduce the greenhouse gas emission and operational expenditure (opex) of data centers infrastructure in the organization. Flexibility design and control of cloud system is backbone of network interconnection facilities.

Gang Sun, Vishal Anand, Dan Liao, Chuan Lu, Xiaoning Zhang, and Ning-Hai Bao[3] describe an efficient power processing through online virtual network. Cloud data centers must attention to pay how much amount of power consumes during data processing for every network request. Its significantly improve utilization factor of clustered nodes. The application request from user could access through This leads to avoid unusual situation of dynamic alteration such as peak load with minimum power utilization effect. Even though, not sufficient to handle huge data centers. It might produce an overheating and wastage of resource inside the data centers.

Anton Beloglazov* and Rajkumar Buyya [5] propose the resource management in virtualized data center method to improve the Quality-of- service and energy utilization. This method doesn't have any knowledge about which application is running in the virtual machine. It's independent of workload type. Require to improve the optimization over the multiple resources and maintain the temperature of cloud centers.

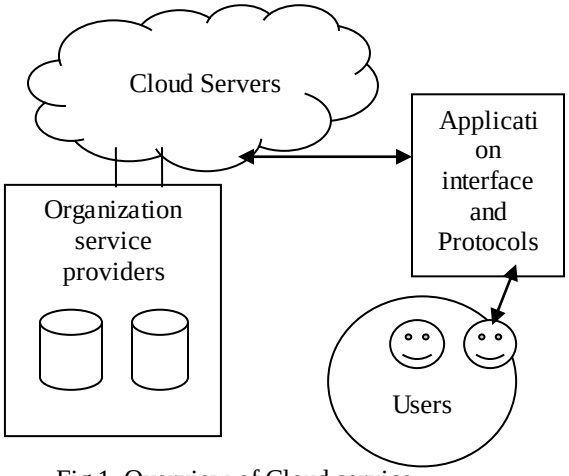

Fig 1. Overview of Cloud service

Above diagram explain the general description of cloud service is accessed by the users. The data centers of cloud server provide services to the N number of users. The virtualized node processes the user request and response through data center.

2.1 Handling of energy consumption

In the cloud computing, data centers consume power in two ways [2] such as network power consumption and server power consumption. Other power consumption is negligible amount and it's connected by the cooling part. If server is turn on, it will utilize certain amount of power is called idle power state. The sum of power consumption is workload-dependent power and ideal power.

$$P(a)=P_{idle}+(P_{busy}-P_{idle})\cdot u$$

virtual machine. This leads to reduce the cost of operational expenditure (opex).
Ahmed Sallam[1,2], Kenli Li[1][4] analyze the scalability of large data handling in virtual machine in cloud environment. To optimize the power consumption, proposed proactive model for an application behavior prediction techniques.

Where P(a) is the total power consumption, P_{idle} is the server running in the idle power state and P_{busy} is the power utilization of all working state. u means utilization in the server.

2.2 Virtual datacenter condition

In the cloud based data center has adjust their data location using some mathematical formulae such as,

$$\text{Minimize } \{\eta\cdot A \text{ network} + \xi\cdot A \text{ server}\}$$

To optimize the total power consumption as the sum of virtual server and network utilization power consumption.

3. Proposed Methods

3.1 Cluster Analysis

In this paper, propose the solution to improve energy utilization factor of data centers in heterogeneous environment. Multiple numbers of users can access the cloud servers time-to-time. They upload or download their data from heterogeneous cloud environment. Data centers can process the request using clustering nodes.

Data centers are maintaining as many service providers like Microsoft, Google, Amazon, and Facebook and so on. They are having all massive cloud servers to sustain their resources and their client resources. The data centers are divided into clustered node. Every clustered node has collection of replicated data. If user request received the data center, it would process their request using active clustered nodes and get their resource. This leads to reduce power consumption and cost of data center.

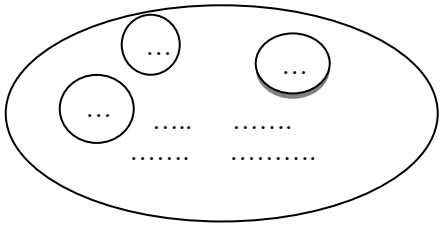

Fig 2. Clustering Nodes

Data can be clustered by using viscovery Somine 6 data mining tool. It's running in the data center (virtual machine server). While client request to the cloud server, the VM server can processing the client request using clustering method. The replicated data accessed

from the active clustered node. These induce low of configuration of clustered node as well as mining data are so fast. Viscovery Somine 6 tool update the clustering nodes and knows how to effectively access the data. This tool maintains the modification of clustering nodes. Based on the observation, clustered nodes are in active sate or in disabled state. Viscovery Somine 6 clustering tool shows the graph of how the data processing happened in the clustered nodes.

Expectation Maximization algorithm

The EM algorithm [9] used to find missing data nodes in the clustering techniques. When the active nodes are arranged based on EM algorithm.

Let $C = \{n^{(1)} \ldots n^{(m)}\}$ be m clustered nodes.
Let $D = \{d^{(1)}, \ldots, d^{(m)}\}$ be m values of hidden data nodes
The clustered node is an observed data given model:
$L(\theta) = \log p(C|\theta) = \log \sum p(C,D|\theta)$
Both D and θ be unknown.
E step: Maximize P with respect to Q, keeping θ fixed.
$Q^{k+1} = p(D|C, \theta^k)$
To estimate distribution over the data of given certain fixed cluster model.
M step: Maximize P with respect to θ, keeping Q fixed
$\theta^{k+1} = \arg\max_\theta P(Q^{k+1}, \theta)$
$= \arg\max_\theta \sum P(D|C, \theta^k) \log p(X, D|\theta)$

Select the fresh arguments for cluster model to maximize expected of observed cluster data and hidden nodes. The algorithm explains the adjustment of clustered data nodes in the cloud server. The active cluster nodes [10] are working based on the frequency of accessibility from database.

3.2 Data Centers

It consists of data warehouse connect with cloud computing service. The data center has N number of cluster nodes to process their operation. IAAS service is maintained within the organization. Use virtualization method to handle more scalable data.

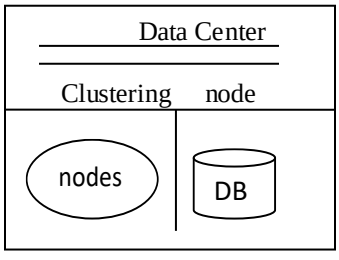

Fig 3. Outlier of data center

3.3 Virtualization

To make data center, we are creating N number of virtual machine using Hyper-V in windows 8.1 PRO since it is easier to implement. Virtual machines are stored in specified drivers. It takes very less space in the system memory. It's in format of image file. The virtual system can handle scalable data effectively using clustering technique.

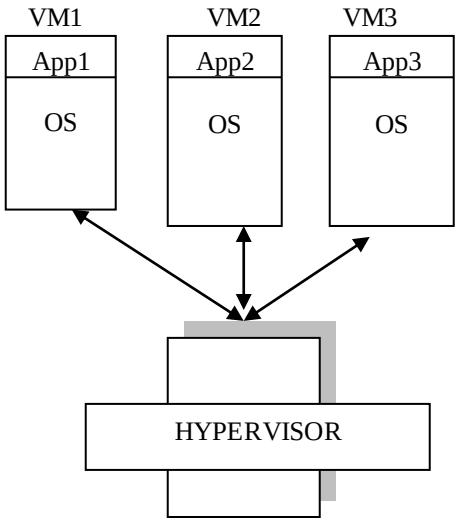

Fig 4. Hypervisor of virtual machine

Hyper-V increases a utilization factor of virtual machine spreading several applications across one set of hardware. Effective Distribution resource allocation and management increased.

4. Result and Discussion

The cloud based data center gain more utilization factor with low power consumption from the organization. Analyze the characteristic of clustered nodes to find an access capability.

Cluster descs	Absolute profile median	Frequency of cluster	Reg. number
C1	0	77.78%	0.143
C2	0	11.11%	0.000
C3	0	11.11%	0.000

Table 1. Description of clustering nodes

The above table explains the active and passive nodes arrangement in the cluster. Here having sample data to derive the result from the cloud server. An application from user is divided into one or more task. After every task submitted in data center, its distribute task equally to the clustered nodes

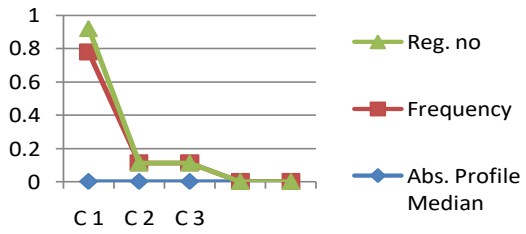

Fig 5.graph of data center accessibility

The graph derived from the sample data processing in the cloud server. In the data center, how the cluster works. The processing capability of each clustered node can be accessed by data center server. The above graph mentioned the frequency of active clustered node. C1. Remaining C2 and C3 are the cluster has less frequent capability to response. So it remains idle until it gets more processing capability. This leads the active node only consume power and have some utilization factor to provide service through the cloud service. We can reduce the maintenance cost of IAAS service. The enormous amount of data is processing effectively in the cloud-based data center.

Based on the comparison with prior work, the heterogeneous energy utilization method is more sufficient to handle the data center.

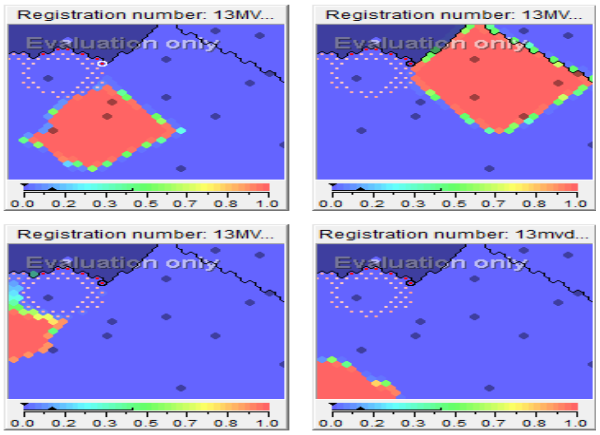

Fig 6. Adjustment of cluster node

The above graph describes the result of clustering node from the viscovery Somine 6 tool. Here have taken some sample worksheet and shown how the clustered node arranged in the system. The data center can consume less power when compare to the prior researchers work. Above the simulation result shown that how much data center can be accessed by user request. So energy utilization factor is minimized.

5. Conclusion

The Heterogeneous energy capacity method shows how the cloud-based data center handle vast amount of data with minimum power utilization factor. Earlier work explains how to handle the workload in heterogeneous environment. In the cloud environment can access different kind of data and their power utilization factor. In this paper describes the maximum workload and minimum utilization factor of energy consum105ption in the cloud server. The virtualization could optimize the infrastructure of cloud service. This method can also sufficient to process multiple data centers in parallel way. In future, will work further to improve the task scheduling delay and maintain dynamic adjustment of data centers to handle scalable servers.

6. Reference

[1] Qi Zhang, Mohamed Faten Zhani, Raouf Boutaba, Joseph L.Hellerstain HARMONY: Dynamic Heterogeneity–Aware Resource Provisioning in the Cloud, 2013 IEEE 33rd International Conference on Distributed Computing Systems, vol:2, page no.510-519.

[2] Burak Kantarci, Senior Member, IEEE, and Hussein T. Mouftah, Life Fellow, IEEE, Minimum Outage Probability Provisioning in an Energy-efficient Cloud Backbone, Globecom-2013 symposium on selected areas in communication, vol-4,page no:2879-2884.

[3] Gang Sun, Vishal Anand, Dan Liao, Chuan Lu, Xiaoning Zhang, and Ning-Hai Bao, presented the Power-Efficient Provisioning for Online Virtual Network Requests in Cloud-Based Data Centers, IEEE SYSTEMS JOURNAL ,vol-2,Jan 2013.

[4] Ahmed Sallam[1,2], Kenli Li[1], Virtual Machine Proactive Scaling in Cloud Systems, 2012 IEEE International Conference on Cluster Computing Workshops, vol-4/11,pgno.97-105, sep-2012

[5] Anton Beloglazov* and Rajkumar Buyya, Energy Efficient Resource Management in Virtualized Cloud DataCenters, 2010 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing,vol-2, pgno.826-831, sep-2010

[6] Srinath Perera, Rajika Kumarasiri, Supun Kamburugamuva, Senaka Fernando, Sanjiva Weerawarana and Paul Fremantle, Cloud Services Gateway: A tool for exposing Private Services to the Public Cloud with fine-grained Control, 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum,vol-4,pgno:2237-2246, June-2012.

[7] Amazon elastic computing cloud. http://aws.amazon .com/ec2

[8] Google cluster data- trace of google workload http://code.google.com/p/googleclusterdata.

[9] Notes for the clustering techniques. mlg.eng.cam.ac.uk/teaching/3f3/1011/lect4.pdf

[10] Gilles Celeux, Gérard Govaert, A Classification EM algorithm for clustering and two stochastic versions, Journal of computation statistics and data analysis,vol-14, issue 3, Oct.1992, pg. 315-332.