

Allocation of Dynamic Virtual Machine using Fuzzy c-means

Lizvita Sharal Tauro*, Gerald Hary Fernandes

Department of Computer Science and Engineering, St. Joseph Engineering College, Mangaluru, India

Abstract Cloud deals with maximizing resource usage in an efficient way. Resource Management and efficient use of that is a biggest challenge. For simulation purpose CloudSim 3.0.3 framework and to display the output in eclipse Luna IDE along with JAVA programming language this research used. Fuzzy c-means is a clustering technique for allocation of dynamic virtual machine in Cloud. Comparison in the CloudSim between the proposed system Fuzzy c-means with the existing system K-means which is one of the clustering technique for allocation of dynamic virtual machine in cloud. The two algorithms had two datacenters, two clusters. Keeping same number of virtual machines and cloudlets in both clustering techniques taking different scenarios, result obtained is that proposed system Fuzzy c-means algorithm took less total execution time in CloudSim than existing system K-means, this means that Fuzzy c-means increases performance.

Keywords Cloud, Clustering, Virtual Machine, Fuzzy c-means, CloudSim

1. Introduction

Cloud Computing provides shared resources, data to devices requested by the user on demand. Cloud deals with maximizing resource usage in an more efficient way. User can access computing resources from shared pool. By paying to the service provider user is able to store their data in own cloud or can make use of third-party datacenter as per requirement of temporary storage. By use of internet, cloud services are reached to customers. Resources can be application based, storage based etc. The services provided by the cloud are subscription based, when users want to use the services they will pay as much they utilize. Datacenters can be located anywhere in the world which has all resources. When user request for any type of resources, nearer datacenter will allocate that can be a virtual machine or any other type of resources if that nearer datacenter has no sufficient resources other datacenter can allocate to user that is not known to the user from where resources are allocated. Resources of cloud not only shared by many users but it has capacity of dynamic reallocation on demand [1]. Sharing of resources in datacenter based on the algorithm is one of the key aspect.

Virtualization technology can be defined as a system acting as a multiple virtual copy so that at same time if many user request for machine each will get access to it. Copies of virtual machines created from physical machines need to

have feel of real machine that can be used a resources. Hypervisor helps to create more virtual machines in the physical machines based on the capacity. Creating virtual machines which executes task has made great benefits in data centers as well in infrastructure models. Nowadays use of cloud has increased, request for resource has been increased scheduling them is also a issue. Resource Management and efficient use of that is also a biggest challenge. If resources are not managed efficient way it directly effects performance of system, cost and indirectly effects functionality of the system.

Simulation which is physically not present but will get a feel how a real cloud works. Before developing a real cloud environment to check utilization of resources alternate to that is by use of simulation tool, deploying and executing by simulation and testing through different test cases. To understand the model how it works simulation is easy way and has less requirement to use. In simulation all concept of cloud can be understood if used for academic purpose. There comes a problem for the provider if he wants to consider realistic workloads which will be large in size and huge data will be present. CloudSim is one of the simulation tool used by the developer for the testing purpose to check performance with no cost which reduces bottleneck before deployed in the real world [1]. A library tool can be used to check different scenarios by executing JAVA code. CloudSim simulator is used to implement different modeling algorithms. CloudSim is a simulator technologies used for abstraction but not any real software runs. Inside the CloudSim framework JAVA code will run. If any one wants to use CloudSim then need to have some basic knowledge about cloud and some basic knowledge about JAVA

* Corresponding author:

lizvitasharaltauro@gmail.com (Lizvita Sharal Tauro)

Published online at <http://journal.sapub.org/ajis>

Copyright © 2017 Scientific & Academic Publishing. All Rights Reserved

programming language. To use this framework and JAVA code requires a software called eclipse which is available in different versions alternate to that there is NetBeans IDE with different version can also be used.

2. Related Work

Abhilasha Singh refers to existing system in which for study purpose from Google, yahoo restricted cloud traces where taken analyzed and categorized workload patterns. Different pattern analysis makes quality better that is user requested for services will be available. Workload Classification on the cloud environment is based on how much cloudlet has consumed resource patterns is done by simulating different parameters of cloud. Intra clustering algorithm was used and had only high level observations, did not contain many variety of workloads and was not able more users and tasks. So proposed a system which could check the user and task behavior patterns using K-means algorithm used for clustering virtual machines in the CloudSim framework. As a result execution time of proposed system was better than the existing system. [1]

Zen Xiao et al. introduces a system which support green computing to effectively use number of servers and also on demand allocation of resources from the data center dynamically for the applications by using virtualization. Algorithm works on to achieve two goals: First is, Overload avoidance which means physical machine need to be capable enough to allocate necessary resources to the virtual machine running on physical hardware else virtual machine may reduce its performance running on physical hardware. Second is, Green computing means if less number of physical machines are capable enough to serve all virtual machines which are created then unused virtual machines can be switched off to save power. These two goals can be balanced by the system autonomous resource management. Predicting resources required by the virtual machine in the future is done by checking the past behaviors of virtual machines. If skewness is reduced after calculation overall use of different resources can be improved. [2]

Rodrigo N. Calheiros et al. Cloud Computing provides different types of services according to industry. Different types of cloud applications example social networking have their different deployment requirements etc. After setting up a real cloud application which includes hardware, software cost, energy consumed, performance may be known is a challenging problem which might have advantage or a disadvantage. So CloudSim framework can be used to check all these criteria with very less cost or no cost by simulation. According to output of simulation then can proceed to set up a real cloud application. [3]

El-Refaey M.A. et al. discuss about workload characterization in the virtual machines. Virtualization has become important in the field of IT now a days as well as in datacenters. Workload characterization is an important aspect that is used for identifying and observing task so that

understanding network is easier. To check the behavior of mail, various application such as editors then Applications can be used. In this paper for classification of workload pattern VCon model is used. This paper discuss about VMware VMMark BenchMark whose final goal is to measure a performance of virtualization with different types of hardware platform across. Server consolidation combines and collects different workloads in a single physical server to ensure all system resources are utilized efficiently. Resources may include network, disk etc. If demands for the resources by the user are balanced then virtualization environment runs smoothly. [4]

Anand M Baswade et al. explain about selecting center point in K-mean algorithm. K-mean belongs to non-overlapping type that is hard clustering, here K which is a object belongs to only one cluster. According to paper one of the strength of K-mean is easier to understand and implementation and one of the weakness of K-mean is it does not handle noisy data. Here traditional K-mean algorithm in which centroids are randomly chosen is compared with new proposed system that is to take initial centroid point by first taking average of n objects. Comparing these two traditional K-mean algorithm and new proposed system that is selection of initial centroid points, new proposed system where initial centroid point is chosen gives more accuracy. According to iterations taken in K-mean algorithm and in initial centroid chosen algorithm K-mean algorithm takes more iteration. So authors say that if center point initially is taken by some modification in the algorithm instead of selecting randomly like in K-mean it can increase performance, accuracy and can reduce number of iterations in the algorithm. [5]

3. Existing System

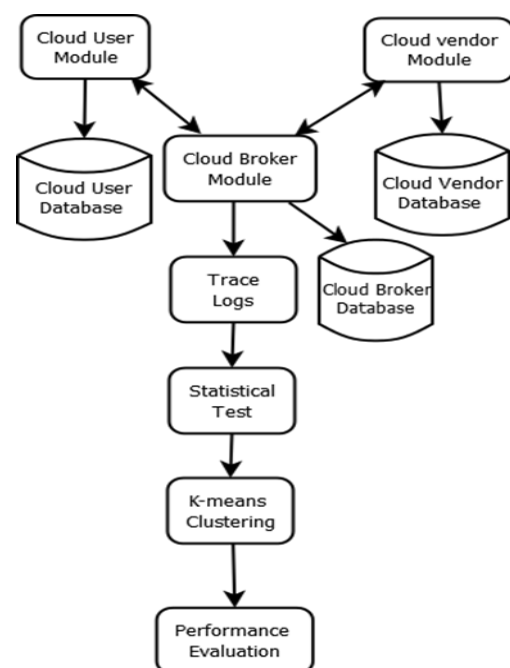


Figure 1. Clustering Approach using K-means

A clustering approach where k-mean clustering algorithm is used to cluster the virtual machines to datacenter.

3.1. Existing System Architecture Design

Existing System uses K-means algorithm, that clusters the virtual machines to groups in datacenter. Figure 1. shows the architectural diagram of existing system. Cloud Broker Module takes the user request to execute task and sends to specific datacenter. Broker Module is intermediate between user and service. Each module have their own databases for storage purpose.

4. Proposed System

A clustering approach where Fuzzy c-mean clustering algorithm is used to cluster the virtual machines to datacenter.

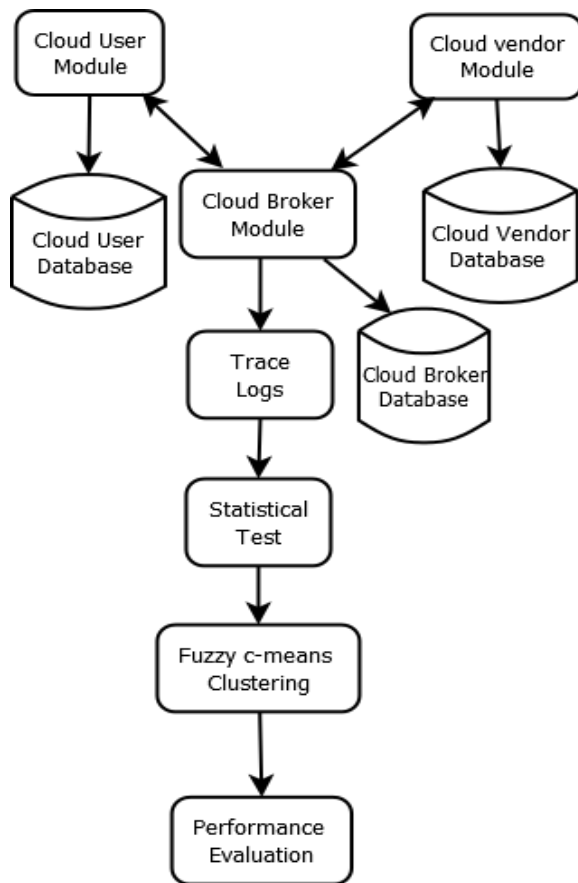


Figure 2. Clustering Approach using Fuzzy c-means

Proposed System uses Fuzzy c-means algorithm, that clusters the virtual machines to groups in datacenter. Figure 2. shows the architectural diagram of proposed system. Cloud Broker Module takes the user request to execute task and sends to specific datacenter. Broker Module is intermediate between user and service. Each module have their own databases for storage purpose. In proposed system Fuzzy c-means technique is used for clustering virtual machines as a group.

5. Clustering using Fuzzy c-means

Fuzzy K-means or Fuzzy c-means is derived from K-means algorithm that is most commonly used and is a simple clustering type. Difference between these two clustering technique is: In K-means also called as hard clusters because one virtual machine or data point cannot belong to two or more clusters where as in Fuzzy c-means virtual machine or data point can belong to two or more clusters with checking condition of probability.

Algorithm:

Virtual machine can belong to only one or more number of clusters. With Fuzzy c-means distance between data points is measured and representation used is n-dimension in the given vector space.

Similar to K-means,

✓ Initialize K clusters

✓ Until intersected

- Calculate the probability of every pair of virtual machines belonging to a cluster.
- Recalculate the center of the clusters by use of previous step of probability membership.

6. Performance Analysis

Figure 3 shows the graphical representation of total execution time of program K-means and Fuzzy c-means clustering technique in CloudSim. Graph clearly shows that Fuzzy c-means takes less time in executing than K-means clustering technique. This graph clearly explains that when less number of virtual machines are present K-means and Fuzzy c-means both are taking almost same execution time. When virtual machine and cloudlet increases Fuzzy c-means takes less execution time than K-means.

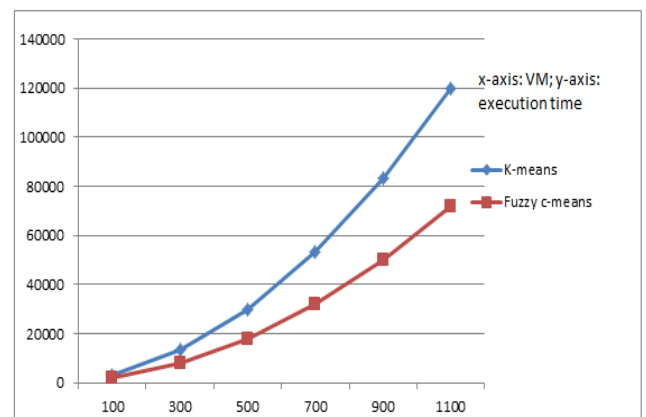


Figure 3. Graphical Representation of total execution time

7. Conclusions and Future Work

The result obtained from research is that proposed system Fuzzy c-means algorithm take less total execution time in CloudSim than existing system K-means. When there are

more cloudlets need to be executed in the virtual machines better to use Fuzzy c-means technique. When there are less number of cloudlets need to be executed in the virtual machines both clustering techniques takes equal execution time. This results indicate that Fuzzy c-means increases performance by executing in less time compared with K-means.

Drawback in Clustering technique used that is K-means and Fuzzy c-means technique while considering uneven request for virtual machines to execute cloudlets that can make unbalance of virtual machine allocation. By different policy this can be made better. Comparing idle time of virtual machines in both K-means and Fuzzy c-means can be considered in the future. Clustering are classified in different types, that can be considered to compare with the proposed system Fuzzy c-means algorithm.

REFERENCES

- [1] Abhilasha Singh 2015. "A Simulation Based Analysis and Modeling of Workload Patterns using the CloudSim Toolkit." (IJSER) International Journal of Scientific Engineering and Research. 36-40.
- [2] Zhen Xiao, Weijia Song, and Qi Chen 2013. "Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment." IEEE Transactions on Parallel and Distributed Systems. 1107-1116.
- [3] CloudSim: A Novel Framework for Modeling and Simulation of Cloud Computing Infrastructures and Services <https://code.google.com/p/cloudsim>.
- [4] El-Refaey, Abu Rizkaa 2009. "Virtual Systems Workload Characterization: An Overview." IEEE Intl. Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises. 72-77.
- [5] Anand M. Baswade, Prakash S. Nalwade 2013. "Selection of Initial Centroids for K-Means Algorithm." (IJCSMC) International Journal of Computer Science and Mobile Computing. 161-164.
- [6] Ranjana C, Theresa Jose 2014. "Dynamic Resource Allocation for the Cloud Users." (IOSR-JCE) International Journal of Computer Engineering. 49-55.
- [7] M. Gavine Kanakadurga Santan B. Veeramallu 2014. "Dynamically Allocating the Resources Using Virtual Machines." IJCSIT 4646-4648.