Imperial Journal of Interdisciplinary Research (IJIR)
Vol-2, Issue-6, 2016
ISSN: 2454-1362, http://www.onlinejournal.in

VM Allocation Using Clustering Technique for Load Balancing In Cloud

Suman B Shahapur[1], Kuldeep P Sambrekar[2]
[1]M. Tech student, Dept of computer science, GIT College, Belagavi, India
[2]Assistant Professor, Dept of computer science, GIT College, Belagavi, India

Abstract: *Cloud computing is an emerging new paradigm that delivers on demand IT services to consumers. Availability, confidentiality, privacy and performance are major concerning areas of cloud computing. VM clustering allows efficient sharing of virtual machines to available datacenters and this clustering technique help to evaluate and enhance the cloud performance. Different allocation policies are available and they have their own advantages and limitations. In this paper a new VM allocation policy is introduced that takes VM`s as per user requirement and allocate them in cluster form to the available datacenters. These clusters of VM`s are formed by using K-Means clustering algorithm. Thus paper in brief discussed about clustering and its importance in service oriented cloud architecture.*

1. Introduction

Load balancing is an approach to reassign the loads from overloaded nodes to underutilize nodes. It is generally dynamic in nature because of traffic flow and need of server, node is depending over the user request. When a node is getting over loaded through user requests then we need a load balancing technique to reassign the loads. The main points to be considered are estimation of load, comparison of load, stability of different system, performance of system, interaction between the nodes, and nature of work to be transferred, selection of nodes. This load considered can be in terms of CPU load, amount of memory used, delay or Network load. There are two types of algorithms used in load balancing techniques 1) static load balancing, 2) dynamic load balancing algorithm. In static load balancing algorithm, fixed no. of steps and prior knowledge is used for load balancing and it cannot depend over the current status of the network whereas in dynamic algorithm load changes according to the current system status of the network. Generally, dynamic algorithm works better then static algorithm. Numbers of load balancing techniques are available can be compared or characterized on following parameters:

- Maximum throughput- Maximize number of finished user requests in a define time unit.
- Completion time- The Maximum time unit required to complete a job.
- Communication cost- The overall cost of transmissions and receiving of the data bits.
- Resource utilization- utilize a resource in such a way that it never get free.
- time- Maximum time required to execute a user request.
- Scalability- future scope to extend the network resource.

2. Load Balancing

Load balancing is the process of reassigning the loads from overloaded nodes to underutilized nodes, to make effectual resource utilization and to recover the response time. Load balancing is one of the major issues connected to cloud computing environment. It is always necessary to share work load amongst the variety of nodes of the disseminated system to get better resource utilization and for better presentation of the system. The goals of load balancing are:

- Improve the performance.
- Preserve system stability.
- Construct fault tolerance.

3. Related Work

3.1 ***A fast adaptive load balancing method***: D. Zhang et al.[1] proposed this method to improve the performance of distributed simulation system. This fast adaptive load balancing method is used to adjusting the workload between the processors from local areas to global areas. In this method a binary tree structure is used to partition the simulation region into sub-domains. Domain decomposition is based on discrete approximation method, finite element method and binary element method.

Advantages of this technique are, the lower communication is overhead, faster balancing speed and high efficiency and the disadvantage is it cannot maintain the topology of cells.

3.2 ***Heat Diffusion Based Dynamic Load Balancing:*** Yunhua deng at el. [2] proposed an algorithm that is based on the principle of heat diffusion, a simple concept is use for load balancing. In heat diffusion concept heat diffusion has happened from high temperature to low temperature. These same phenomena is used for load balancing purpose in the terms of VM's the traffic flow of user request is from overloaded VM to under loaded VM. According to this algorithm the virtual environment is divided into number of cells and each cell have objects, every node in cell send load information to its neighbour node in single iteration.

In heat diffusion load balance environment amount of load migrates is minimum. This algorithm Efficient for multiprocessor network and network latency is minimizing for load transfer between the cells. Limitations associated with this technique are very small connectivity for large scale graphs or cells, higher computational and communication methods are used, network delay on single path, and if there are more iteration then more time wastage.

3.3 ***A Dynamic and Adaptive Load Balancing Strategy For Parallel File System:*** Bin dong et al. [3] proposed a technique where it is based on distributed architecture for dynamic and adaptive load balancing. They proposed an algorithm called as a SALB (self active load balancing algorithm).In parallel file system data is collectively transferred between memory and I/O systems in a single transmission. Using this technique data is effectively transferred over a cloud structure and load balancing is needed for dynamic file migration.

Advantages of this technique are high speed processing, elasticity, based on global load management, resource utilization, works on large file system, and load migration without affecting to system processing. A limitation associated with this algorithm is that there are some file migration side effects that degrade the system performance.

3.4 ***Honey bee behaviour Inspired Load Balancing:*** Dhinesh et al. [4] proposed an algorithm named honeybee behaviour inspired load balancing algorithm. In HBB-LB technique same idea is used for load balancing, the tasks are denoted as honey bees and VM's are denoted as food sources. When a VM is find under loaded then it like foraging bee finding a new food source. Then a status has indicated like perform a waggle dance to show that how many tasks can be perform the VM and task can be chosen according to their load and priorities. This updating status will give the clear idea about which task is assign to which VM based on the availability and scalability and total load of VM.
HBB-LB proposed algorithm for minimizing the calculation of current workload, good resource utilization, high throughput, and QoS is based on the task priority. Limitations related to this algorithm are that it may suffer from load imbalance between the VM's, minimizes the response time and underutilization of low priority task.

3.5 ***Shared Resource Clustering for Load Balancing:*** Dr. Vinay Chavan et al. [5] addressed this technique for load balancing in cloud. In this method jobs are made up of different tasks; Virtual Machines (VM) are required to execute this set of tasks. Clustering these tasks improves the level of performances by balancing load and on less number of VM, which are required for executing single task at the same time. If task are not clustered then single jobs are assigned to virtual machines and need of deploying these virtual machines with dynamic creation increases in real time. In addition we calculate deadline constraints for user tags while doing task clustering.

4. Proposed Clustering Technique for Load Balancing

4.1 Proposed system:

These jobs are made up of different tasks; Virtual Machines (VM) are required to execute this set of tasks. Clustering these tasks improves the level of performances by balancing load and on less number of VM, which are required for executing single task at the same time. If task are not clustered then single jobs are assigned to virtual machines and need of deploying these virtual machines with dynamic creation increases in real time. In addition we calculate deadline constraints for user tags while doing task clustering.

Task clustering results in creating lesser number of jobs and improves the load balancing mechanism in cloud environment, to distribute workload and balances the resource which will help in increasing the performance and network flow in distributed environment in cloud computing.

4.2 System design

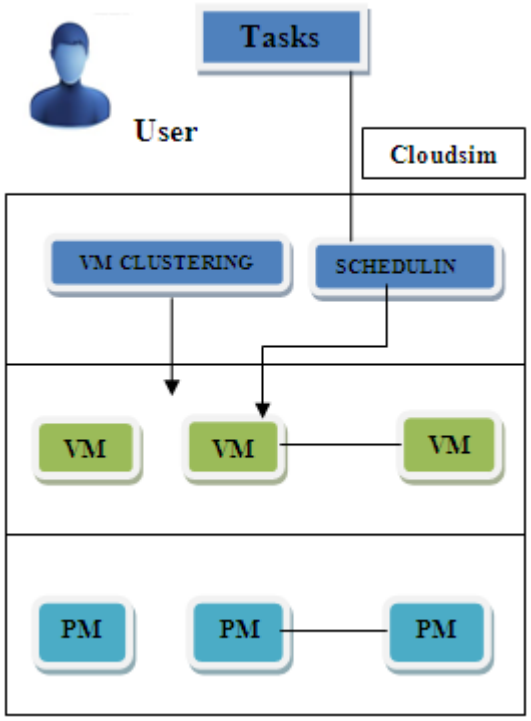

Figure1. System design

Mainly Clustering is the method which includes the grouping of similar type objects into one cluster and a cluster which includes the objects of data set is chosen in order to minimize some measure of dissimilarity. For scheduling the virtual machines, K-Means clustering algorithm is used. K-Means clustering is a clustering method in which the given data set is divided into K number of clusters. K-means clustering is a well known partitioning method. In this objects are classified as belonging to one of K-groups. The results of partitioning method are a set of K clusters, each object of data set belonging to one cluster. In each cluster there may be a centroid or a cluster representative.

Clustering algorithm: K-Means follows the partitioned or non-hierarchical clustering approach. It involves partitioning the given data set into specific number groups called Clusters. Each cluster is associated with a centre point called centroid. Each point is assigned to a cluster with the closest centroid.

Implementation of proposed technique is based on cloud based simulation, using Cloudsim. This simulation supports in the distributed environment and techniques for implementing cloud with open access programming for modelling and simulating cloud computing infrastructures and services supports for the execution of workloads.

5. Conclusion

This proposed work was aimed to improve the performance of cloud computing with VM clustering. It arranges virtual machines in cluster form before allocating them to the datacenters. This arrangement provides efficient CPU utilization and load sharing among the datacenters, so performance can be enhanced in some aspects. This work is performed to analysing performance parameters and its enhancement using clustering technology. Although integration of K-Means clustering works well than the existing methodologies in cloud environment but this also can be replaced with other available clustering techniques.

6. References

[1] Dongliang Zhang, Changjun Jiang,Shu Li, "A fast adaptive load balancing method for parallel particle-based simulations", Simulation Modelling Practice and Theory 17 (2009) 1032–1042.

[2] Yunhua Deng, Rynson W.H. Lau, "Heat diffusion based dynamic load balancing for distributed virtual environments", in: Proceedings of the17th ACM Symposium on Virtual Reality Software and Technology, ACM, 2010, pp. 203–210.

[3] Bin Dong, Xiuqiao Li, Qimeng Wu, Limin Xiao, Li Ruan, "A dynamic and adaptive load balancing strategy for parallel file system with large-scale I/O servers", J. Parallel Distribution Computing. 72 (2012) 1254–1268.

[4] Dhinesh Babu L.D, P. VenkataKrishna, "Honey bee behaviour inspired load balancing of tasks in cloud computing environments", Applied Soft Computing 13 (2013) 2292–2303.

[5] Dr. Vinay Chavan, Parag Ravikant Kaveri "Clustered Virtual Machines for Higher Availability of Resources with Improved Scalability in Cloud Computing" Networks & Soft Computing (ICNSC), 2014 First International Conference IEEE.

[6] Bhupendra Panchal and R. K. Kapoor, "Performance Enhancement of Cloud Computing with Clustering", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-5, June 2013.

[7] Youseff, L. Toward a Unified Ontology of Cloud Computing Retrieved from http://spsteve

[8] Prakashan Korambath, Ilkay Altintas "A Physical and Virtual Compute Cluster Resource Load Balancing Approach to Data-Parallel

Scientific Workflow Scheduling" Jianwu Wang1, 2011 IEEE World Congress on Services.

[9] Load Balancing in Cloud computing, http://community.citrix.com/display/cdn/Load+Balancing .