Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 3, July 2021:2951- 2964

Research Article

Resource Optimization and Dynamic Workload Prediction Strategy for Large Scale Cloud Environment

S.Kalaiselvi [1], R.S.Mohana [2], P.Prabakaran[3]

Abstract

Cloud computing is a ubiquitous computing paradigm which eliminates the technical obstacles which the organizations have to deal with and provides IT services on-demand. Cloud computing is a cost-effective model for provisioning services and makes the ICT management effective in dealing with real-world workloads. It is becoming popular since it provides better usability at a lower cost leading to higher utilization and better management. Clustering based approach is used for the allocation of virtual machine to the incoming request. Here in the proposed system a hybrid approach is used with kernel based fuzzy C-means clustering mechanism with Bayesian prediction for the provisioning of resources in the heterogeneous cloud environment. The results are optimized with particle swarm optimization. Hence the total average waiting time is minimized and the results are evaluated using Cloudsim toolkit and the results reveals that the hybrid system achieves better performance when compared to the existing technology.

Keywords: *fuzzy clustering, resource scheduling, workload prediction, PSO, bayesian.*

Introduction

[1]Assistant Professor(SLG).Dr, Kongu Engineering College, Department of Computer Technology-UG, kalai.ctug@kongu.edu

[2]Associate Professor.Dr, Kongu Engineering College, Department of Computer Science and Engineering, mohanapragash.cse@kongu.edu

[3]Assistant Professor(SRG)., KPR Institute of Engineering and Technology, Department of Computer Science and engineering, prabakaranmp66@yahoo.com

Received: 01.01.2016, Accepted: 01.01.2016

Cloud computing is a paradigm that serves an off-promise computing of Internet. This technology offers proficient computing by consolidating bandwidth, processing and storage memory. The IT services is expected to concentrate further on optimizing the information design by optimizing the information design by using the on-demand IT services instead of interior performances and infrastructure which is not necessarily known to the user. The key component of cloud computing is virtualization, which provides a different environment to the user with specific customization features. Cloud services have three distinct characteristics which differentiate them from traditional computing. It is sold on demand, elastic and the service is entirely managed by the provider.

Virtual provisioning is intended to reduce storage management by permitting storage managers to meet the demand for capacity on-query. The concern for virtual provisioning is that it needs managers to supervise the handling of virtually provisioned resources with effective methodologies. Grouping the similar resources improves the overall resource management performance in cloud environment. In order to overcome the drawbacks with hard clustering technique, a soft clustering has been introduced. A soft clustering method provides promising performance based on the flexible nature. Fuzzy C-means (FCM) is one of the most important technologies for spherical clusters. For general clusters the kernel-based technique with higher-dimensional feature space is used in the proposed system. Here it is very much important to select the best kernel among the extensive range of possibilities.

The main aim of the system is to suggest an appropriate clustering-based task scheduling model for cloud computing because it cluster the virtual machines according to the MIPS, RAM, SIZE of virtual machines. The proposed work is to cluster the virtual machines using kernel based Fuzzy C means clustering algorithm.

KFCM models the uncertainty and ambiguity of data and ensures the initial knowledge of the application behavior in the prediction models. The rules, membership values and inference systems should be chosen consciously to obtain accurate results. Prediction methods are applied to each cluster and object. Hence the time and cost of prediction get reduced and the correlation is inferred among the available resources, thereby improving the prediction results.

Related work

Bonomi and Kumar [6] proposed a optimal load balancing strategy based on the effective resource scheduling algorithm in cloud computing. The performance evaluations are done with respect to load-balancing, computational efficiency. This approach handles a heterogeneous environment in a stable state without including the dynamic changes in communication aspects.

Shyamala and Sunitha Rani [1] have proposed the resource allocation methodologies for balancing the load of the available resources so that the resource utilization is improved. Piotr Nawrocki, Wojciech Reszelewski [20] have presented the different scheduling approach for reducing the delay time. This proves that extending the common cloud infrastructure policies to the Mobile cloud computing environment effectively minimizing the resource demand.

Shinde et al [21] proposed the task scheduling using Fuzzy C-Means (FCM) algorithm as a clustering technique and Linear Programming approach as optimization technique. The goal is to allocate the user tasks to the cloud resources in such a way that it minimizes the cost that user is required to pay to the cloud owner without exceeding execution time limit. The results of proposed strategy are compared with the results observed by First Come First Served (FCFS) scheduling. With the obtained results using linear programming approach as optimization along with FCM as clustering methodology, so that the objective of cost and resource optimization is achieved.

The video segmentation technique [26] using FCM clustering which denotes the existing approaches used in video segmentation are analyzed. reviewed that FCM based video segmentation algorithm are commonly used in recent image and video processing applications. Kernel based mechanism used for adjusting the weights automatically and provides effective output for improving the overall performance of cloud. Venu and Anuradha [13] presented FCM based technique in analyzing the image segmentation. Hybrid hyperbolic tangent kernels and Gaussian kernels are utilized for clustering the available set of

images. The performance is tested by means of score, number of iterations and execution time under the different Gaussian noises.

In the task based load balancing model [11] with particle swarm optimization, the QoS in cloud improves by eliminating the delay in the virtual machine. In deadline based resource scheduling [15], a heuristic optimization approach with PSO technique for scientific workflow on IaaS clouds enhances the quality of service.

Workflow pattern prediction for cloud services differs from one application to another. Zhang et al [20] presented deep learning model for workload prediction of cloud computing. Jitendra Kumar, and Ashutosh Kumar Singh [21] applied artificial neural network approach for effective work load prediction. Sahi and Dhaka [24] states some effort on workload prediction and load management. For instance Rodrigo et al [22] introduced a workload prediction algorithm based on ARIMA model. The workflow analysis mainly concentrates the load balancing techniques to adjust the proper scheduling of resource and thus balance the capacity.

Clustering

Cloud computing is a new technology. It provides online resources and online storage to the user's .It provides all the data at a lower cost. With the advent of cloud computing the resources can be accessed from anywhere at any time with the help of Internet. They need to pay for what they have used. Clustering is the mechanism of grouping the data points into independent groups or clusters so that items in the same cluster are dependent with one another as possible and items in different classes are independent with each other.

Clustering is also used for data compression, in which large amount of data are converted into less number of models or clusters. It comes under the category of unsupervised learning without any training and testing phase instead the clustering is done by observation. The cluster should exhibit two main properties; low inter-class similarity and high intra-class similarity.

Fuzzy c-mean algorithm is used for clustering the virtual machines on the basis of some parameter. The VM that are clustered do complete belong to one cluster but they have their membership in each cluster. Based on the membership value the virtual machines belong to the cluster to which it has its maximum membership. Fuzzy c-means is an effective soft clustering mechanism belonging to an unsupervised method for the analysis of data and construction of models. In many scenarios, fuzzy clustering is more applicable when compare to other clustering. Objects lies on the boundaries between several classes are not forced to become the member of one of the classes, but rather are assigned membership degrees between 0 and 1 indicating their partial membership. Fuzzy C means clustering was applied in many research problems in which the membership value lies between 0 and 1.

Fuzzy c-means algorithm is most widely used technique that employs fuzzy partitioning such that a data point can belong to all groups with different membership values between 0 and 1. The fuzzy c-means algorithm used to classify the given set of elements into the collection of clusters.

Proposed work

Resource provisioning has been an important problem in the cloud environment. To solve this problem the proposed system employs a hybrid mechanism with Fuzzy C means and improves resource utilization and provisioning in minimum time and to meet Service Level Agreement (SLA). In FCM where each item belong to more than one group where degree of membership for each item is given by a probability distribution over the clusters. It is useful when required number of clusters are pre-determined. The concept behind the fuzzy clustering is the variable length partitioning of the data into a group of clusters. Its advantages are unsupervised and better convergence.

Resource provisioning plays a vital role in the cloud environment. To handle this issue in cloud, resource utilization and provisioning in minimum time to meet the Service Level Agreement (SLA). Particle Swarm Optimization (PSO) is a most widely used optimization

technique based on the social behavior of birds. It is the commonly used technique for effective resource provisioning. Particle Swarm Optimization (PSO) have been found to be very useful in optimization area.

Generally PSO, based on the swarm of particles and it is the population based optimization algorithm. Here the position of the particles may be changed in the given search space based on the own experience and its neighbor based on its velocity. The particles fly in the high dimensional space based on the current optimum particles. PSO is initialized using the random solutions and then searches for optimal value in each iteration. In every step, the particle is changed by following the best solution it has achieved known as stored fitness value.

When an independent particle in the population as its immediatel neighbors, the best value is known as ilbest. After determining the two consecutive best values, the particle updates its velocity and positions with equation (1) and equation(2).

$$v_{i,j} = x * v_{i,j} + c_1 * r_{1\,Y} * (p_{p,j} - q_{i.j}) + c_2 * r_2 * (p_{g.j} - q_{i.j}) \quad (1)$$

$$p_{i.j} = p_{i.j+} vi. \quad (2)$$

In equation 1, ‘x’ is the inertia factor which influences the local and global settings of the algorithm. , qi.j is the velocity of the particle ‘I’ ,and c1 and c2 are weights of the cognitive and social factors. r1and r2 (0,1). pp.j denotes the best value determined by particle ‘i’ (pbest) and pg.j gives the global best value computed by the entire swarm (gbest).

The proposed work aims in clustering the virtual machines based on the multiple kernel based FCM approach optimized with PSO optimization technique. FCM determines the number of clusters and virtual machine parameters such as MIPS, RAM and VMsize. Generate the random cluster center values based on parameter values.

Virtual machines are added to the appropriate clusters based on the workload prediction with Bayesian model. The computational intensive tasks are served based on the hybrid approach to improve the user satisfaction. Hence the quality of service is maintained and the tasks are

served without violating service level agreement. The architecture of the proposed system includes the cluster controller for effective clustering approach and optimized with PSO optimization technique.

Here FCM is used to find the final centroid values and cluster the VMs based on the small, medium and large categories. Finally virtual machines are clustered using FCM. Next PSO is implemented where virtual machines and jobs are implemented as row and columns respectively in execution, population matrices. Generate execution time for each machines to do a particular tasks. Allocate the task to the corresponding VMs based on minimum execution time. In population matrix generate machine numbers uniquely. Tasks are allocated to each clusters efficiently. Thus makespan is better by combining FCM and PSO.

FCM- PSO-Bayesian Model

Initially the numbers of virtual machines are given to the fuzzy clustering as input then the virtual machines are clustered into three clusters, then Bayesian model is applied for prediction. Based on the predicted values virtual machines are added to the clusters. Fuzzy clustering is a powerful unsupervised method for the analysis of data and construction of models. Objects belonging to more than one classes are assigned membership degrees ranges between 0 and 1 indicating based on the nature of participation in that class. Kernel based FCM handles the non linear data items and capable of solving the real world problems more effectively.

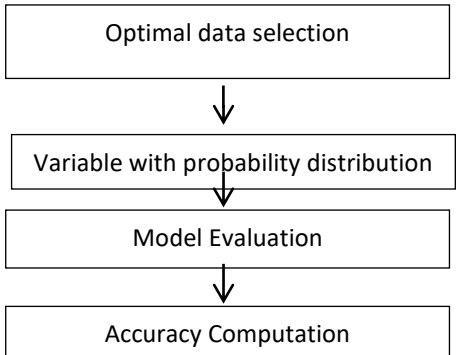

Figure 1. Flow chart of FCM- PSO-Bayesian Model

Bayesian model is commonly used model for capturing the workload pattern. Learning new values based on the utilization of available VMs and response times. These features are used to make the mechanism works well with the given range of data values. This model is applied to predict the future pattern based on the available statistics and analysis is done to determine the actual and observed values, which represents the pattern of CPU utilization at the specific time k and the virtual machines are clustered.

In Bayesian inference, the parameter learning is done with Bayes theorem. The probability function by three variables is defined by

$$P(S1,S2,S3) = P(S2|S1)P(S3|S1)P(S1) \quad (3)$$

In this theorem the prior probabilities are used to compute the posterior probabilities. In general, Bayesian network utilizes the total probability law to compute the joint probability. The likelihood of VMs utility for virtual machines are calculated by applying the chain rule based on the product of independent probability values. The FCM employs fuzzy clustering such that a data point can belong to more than one cluster with varying membership values ranging from 0 and 1. Here the kernel based FCM addresses the non linear nature of real world scenarios when combined with the prediction approach the future demand is addressed. In order to minimize the overall error the non linear optimization approach is employed in the proposed research.

Thus the hybrid approach improves the scalability and reduces the error in resource provisioning of cloud environment. The system attempts to partition a finite collection of elements X={x1,x2,....,xn} into a collection of Fuzzy clusters with respect to some given criterion.

In the experiment, the Virtual Machines are grouped into three clusters as small, medium and large based on different parameters such as RAM size, Million Instructions Per Second(MIPS), and Size of the virtual machine. The cluster is initially assigned with centroid values then the distance matrix and membership matrix are calculated. Total membership value for the value in both sample or decision space must be equal to 1. Then the centroid values are calculated. The process is repeated until the centroid values equal the previous iteration's centroid values.

In this simulation, 512 cloudlets and 20 virtual machines are used for fuzzy clustering and VMs are clustered. Based on the CPU usage values of clusters Bayesian model is used to predict the next week's CPU usage and virtual machines are dynamically created and added to the clusters. Thus it improves the efficiency of the system. Virtual machines are created dynamically based on the predicted values of CPU usage. Virtual machines are added to the clusters dynamically. Cloudlets are binded to the virtual machines created. After completing dynamic VM creation the makespan and total execution time is noted to be reduced compared to the makespan and total execution time calculated after fuzzy clustering.

Results and Discussion

Setting up cloud Environment

To demonstrate the effectiveness of the system, experiment environment is build with 10 Host,20 VMs and 512 cloudlets are created and the server is XenServer. Datacenter is created and create DataBroker for the DataCenter. Then virtual machines and cloudlets are submitted to the broker.Create the array of virtual machines list, each vmlist contain the following parameter vmid, mips, size, ramsize, bw, PostNumber and server name. Let 'n' number of tasks are considered in the proposed system denoted as{T1, T2, T3,…, Tn}. Number of VM is denoted as 'm' {VM1,VM2, VM3, …, VMm}. In this, the tasks are not related with one another. Each VM would execute one task at a time, and the remaining tasks would wait in the queue.Create the virtual machine and add the VM to the vmlist by mlist.add() function. Assign or map the broker Id into cloudlet in order to map the virtual machine Id to cloudlet.

The Parameters used for evaluation are

- Makespan
- Total Execution time

Comparison of Makespan

Makespan- The execution time for each virtual machine is calculated and the maximum value among them is taken as makespan.

$$\text{Makespan}=\max\{CT_{ij} / i \in T, i=1,2,\ldots n \text{ and } j \in VM, j=1,2,\ldots m\} max\{CT_{ij}/i\} \quad (4)$$

The makespan values are compared with various numbers of cloudlets. The combination of KFCM and the PSO algorithm gives optimal solution when compare to the existing scheduling algorithm. Under the normal workload pattern the kernel based FCM behaves like the existing algorithm but with the optimization technique the result shows improvement. From Figure 1 the following conclusions can be made. First the clustering algorithm is proved more effective than the default scheduler and resource optimization improves the performance of the servers, because the smaller the value of the makespan the better the performance of the server.

It can be seen that the scheduling scheme of the proposed algorithm with optimization are more reasonable when combined with the effective clustering approach. It can be seen from the figure that the proposed algorithm adaptively schedules the tasks on suitable servers and produces the significant improvement in maintaining the quality of service under different circumstances.

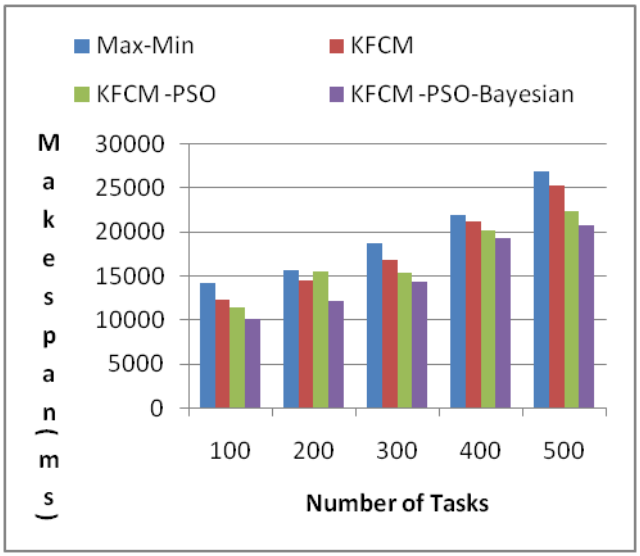

Figure 2. Makespan obtained after workload prediction

Based on the growing need for SaaS based applications on the cloud when compare to the traditional desktop based applications the dynamicity of the resource requirement should not reduce the performance of the server and also affects the user experience. The prediction based on Bayesian model combined with the clustering gives better performance. Here the cluster controller plays a vital role with predictor component to improve the utilization without any wastage.

It is evident from the simulation results that the hybrid optimization with prediction shows a commendable performance improvement in handling the resource management issues of virtualized cloud computing environment. Fig. 3 shows summarizes the comparison of the proposed techniques and it can be seen that the makespan of the virtual machines after the clustering, optimization and prediction techniques. Here the hybrid approach helps to reduce the execution time of independent VMs and hence the overall performance improved.

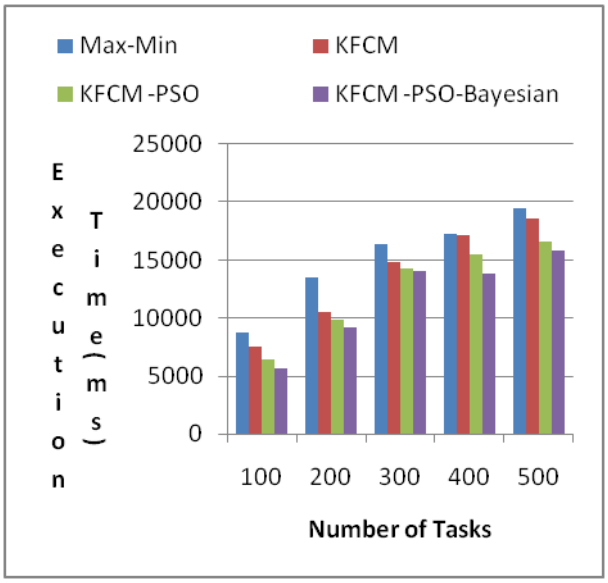

Figure 3. Comparison of Execution Time

With the effective resource provisioning mechanism a cloud environment can meet the service level agreement with the customers.

Conclusion

Thus the proposed system shows how clustering could be done efficiently using Fuzzy C-Mean clustering and Bayesian model is used for prediction of CPU utilization based on past history of utilization. Based on the predicted values virtual machines could be added to the clusters. The proposed architecture aims at achieving the minimum makespan, minimum total execution time and maximum utilization of the cloud resources. The results show that the proposed system reduces the execution time by reducing the makespan when compared to the execution time of non-clustered virtual machines. In the future, focus will be on finding more effective clustering algorithm and prediction model to enhance the performance of the proposed algorithm and the work would also be extended to implement in real cloud.

References

Kesavaraja, D., & Shenbagavalli, A. (2018). *QoE enhancement in cloud virtual machine allocation using Eagle strategy of hybrid krill herd optimization: Parallel and Distributed Computing.* Journal of parallel and distributed Computing , 118, 267-279.

Daniel Graves., & Witold Pedrycz. (2010) *Kernel-based fuzzy clustering and fuzzy clustering A comparative experimental study.Fuzzy Sets and Systems*, 16, 522-543.

Ejaz Ahmed., AbdullahGani., MehdiSookhak., SitiHafizah AbHamid., FengXia. (2015) *Application optimization in mobile cloud computing: Motivation, taxonomies and open challenges*. Journal of Network and Computer Applications, 52, 52-68.

Bonomi, F., & Kumar, A. (2016) , *Adaptive Optimal Load-Balancing in a Heterogeneous Multi-server System with a Central Job Scheduler*, IEEE Trans. Computers, 39(10), 1232-1250.

Fahimeh Ramezani Jie Lu & Farookh Khadeer Hussain. (2014) *Task-Based System Load Balancing in Cloud Computing Using Particle Swarm Optimization.* International Journal of Parallel Programming, 42, 739-754.

Fereshteh Sheikholeslami., & Nima Jafari Navimipour. (2017) *Service allocation in the cloud environments using multi-objective particle swarm optimization algorithm based on crowding distance.* Swarm and Evolutionary Computation, 35,53–64.

Golchi, MM., Saraeian, S., & Heydari, M .(2019). *A hybrid of firefly and improved particle swarm optimization algorithms for load balancing in cloud environments: Performance evaluation*, Computer Networks, 162,1060-1068.

Jitendra Kumar., & Ashutosh Kumar Singh .(2018). *Workload prediction in cloud using artificial neural network and adaptive differential evolution.* Future Generation Computer Systems, 81,41-52

Kalaiselvi, S., & Kanimozhiselvi, C.S., (2020). *Hybrid cloud resource provisioning (HCRP) algorithm for optimal resource allocation using MKFCM and bat algorithm*. Wireless Personal Communications, 111(2), 1171-1185.

Mary Shanthi Rani., Chitra, P ., & G Shanthi, G.(2017). A *Study on video Segmentation using Fuzzy C Means technique*, National conference on Computational Methods, Communication Techniques and Informatics, 1.

Mihaela-Andreea Vasile. , Florin Pop., Radu-Ioan Tutueanu., Valentin Cristea., &Joanna Kolodziej. (2015) *Resource-Aware Hybrid Scheduling Algorithm in Heterogeneous Distributed Computing*. Future Generation Computer Systems,51,61-71.

Nookala Venu., & Anuradha. (2015) *Two Different Multi-Kernels for Fuzzy C-means Algorithm for Medical Image Segmentation*. International Journal of Engineering Trends and Technology 20:77.

Rodrigo N. Calheiros., Enayat Masoumi., Rajiv Ranjan & Rajkumar Buyya.(2015), *Workload Prediction Using ARIMA Model and Its Impact on Cloud Applications*. QoS IEEE transactions on cloud computing, 3

Shabeera, TP., Kumar, SM., Salam, SM., & Krishnan, KM.(2017), *Optimizing VM allocation and data placement for data-intensive applications in cloud using ACO metaheuristic algorithm.* Engineering Science and Technology,an International Journal, 20(2),616-628.

Shinde, MS., & Kadam, A.(2015), *Cloud Based Task Scheduling Using Fuzzy C-Means and Linear Programming Approach,* International Journal of Science and Research, vol. 04, pp. 2319-7064.

Shyamala, K., & Sunitha Rani, T. (2015). *An Analysis on Efficient Resource Allocation Mechanisms in Cloud Computing*. Indian Journal of Science and Technology, Vol 8: 814–821.

Supreet Kaur Sahi & Dhaka, V.S .(2015) , *A Review on Workload Prediction of Cloud Services.*International Journal of Computer Applications,109 (9), 0975 – 8887.

Zhang, Q, Yang, LT, Yan, Z, Chen, Z & Li, P .(2018), *An efficient deep learning model to predict cloud workload for industry informatics*, IEEE transactions on industrial informatics, 14(7), 3170-3178.