International Innovative Research Journal of Engineering and Technology
ISSN: 2456-1983 Vol: 5 Issue: 3 March 2020

Heterogeneous Resource Provisioning In Cloud

M.S. Anbarasi[1], V. Jayasri[2]

[1]Assistant Professor, [2]Master of Technology,
[1,2]Information Technology, Pondicherry Engineering College, Puducherry
[1]anbarasims@pec.edu, [2]jayasri211995@pec.edu

Abstract: Cloud computing is the recent large-scale distributed computing technology, which also allocates resources to the users on demand in cost-effective manner. Resource Allocation (RA) is a dynamic access and process of the cloud resources by the end users who act as a client for Resource Provisioning. Though resource provisioning is a dynamic management technology, yet there exist issues such as resource over-provisioning and resource under-provisioning. The over-provisioning can cost user's more than necessary resource and under-provisioning affects the application performance for the user. Therefore, the efficient usage of resources in cloud for large scale task processing is one of the most challenging tasks. The main aim of this project is to identify optimal task using Ant Colony Optimization (ACO) as first phase then for prioritizing task followed by Analytical Hierarchical Process (AHP) for heterogeneous workload allocation.

Keywords: Resource Provisioning, Resource Allocation (RA), Analytical Hierarchical Process(AHP), Ant Colony Optimization(ACO)

1. INTRODUCTION

In current real-time scenario, cloud computing environment provides computing (C) and non-computing (NC) resources to the Government and private organization based on pay-for-utility for end users. Cloud computing enables cloud users to utilize resources without installation or deployment. In the cloud computing environment resource provisioning problem can be viewed from different perspectives such as, the Infrastructure as a Service (IaaS) Provider, the Software as a Service (SaaS) provider, and the cloud end-user. The provisioning of resources in cloud computing must be managed by the end user, which is a new and unfamiliar concept for the users who are used to work with a fixed set of resources. The end-user is an individual or an organization with the aim of renting computational resources from a public cloud provider. The resource provisioning issue in the cloud computing world needs to be addressed separately for each stakeholder due to different objectives, specifications and constraints. Resource allocation costs per unit are usually the lowest, but under provisioning may occur if the allocated resources are unable to meet the demand in full. To manage all these problems and optimize cost, a complicated decision making problem is used to choose the most suitable type and number of VMs with the best pricing plans, for running an application.

Static provisioning, user contracts with the service provider and the provider will prepare the appropriate resources before the service begins. The client is paid a flat fee or is paid on a month to month premise. Dynamic Provisioning methods have been recommended whereby Virtual Machines (VMs) might be relocated on-the-fly to new process hubs inside the cloud. The provider allocates more resources as required through dynamic provisioning and expels them when they are not needed. The client is charged on a compensation for every utilization premise. To create a hybrid cloud environment dynamic provisioning is utilized, which is sometimes referred as cloud bursting.

The preeminent goal of the cloud user is to reduce the cost by renting the resources and from the cloud service provider's point of view is to maximize profit by proficiently allotting the resources. In order to accomplish the objective the cloud user needs to demand cloud service provider to make an arrangement for the resources either statically or dynamically. So, that the cloud service providers will know what number of occurrences of the resources and what resources are required for a specific application. By provisioning the resources, the QoS parameters like response time, availability, performance etc., must be achieved without violating SLA.

2. LITERATURE SURVEY

For cloud computing, cloud providers provide two provisioning strategies to be specific reservation and on-demand models for cloud customers to manage resources. Generally, the cost of using the reservation system-provided storage tool is less costly than that based by the on-demand service, as the cloud user needs to pay the provider in advance. The user will reduce the cost of all-out resource provisioning with the reservation plan. Be that as it may, it is difficult to achieve the best early reservation of services because of the uncertainty of the future interest of the customer and the resource costs of the providers. By specifying a stochastic programming model, an optimal cloud resource provisioning (OCRP) algorithm is proposed to address this issue. The OCRP calculation can arrangement figuring resources for being utilized in various provisioning stages just as a long haul arrangement, e.g., Four phases of a quarter plan and twelve phases of an annual arrangement. The vulnerability of interest and value is considered in OCRP [1].

Cloud computing provides a dynamic portion of resources on demand, the component that allows it to stand apart in order to provide exceptional execution, flexibility, cost-effective and less maintenance, and then settle a suitable decision on it.The fundamental factor of task scheduling is to do dynamic allotment of assets to build execution and lessening the expense. In this work, an answer is proposed utilizing make-span and cost, accepting them as significant limitations for the enhancement issue. In particular, we consolidated two algorithms, the cuckoo search algorithm (CSA) and the opposition based learning algorithm (OBL), and made another hybrid algorithm called the opposition cuckoo search calculation (OCSA).OCSA algorithm **[2]** indicated observable improvement over the other undertaking planning calculations. The proposed work is recreated in Cloud-Sim programming condition and the reproduction results demonstrate the viability of the proposed work by limiting expense and make-span parameters.

Cloud computing has turned into many new innovation that has tremendous possibilities in agreements. Clouds can make it conceivable to get to applications and related information from anyplace. Organizations can lease resources from cloud for capacity and other computational purposes with the goal that their foundation cost can be diminished essentially. Further they can utilize broad access to applications, in light of pay as-you-go model. There is no need to obtain licenses for individual items. Anyway one of the significant entanglements in cloud computing is identified with advancing the resource being dispensed. In light of the uniqueness of the model, resource portion is performed with the target of limiting the expenses related with it. Different difficulties of resource allotment are fulfilling client needs and application prerequisites. In this paper, different resource allotment methodologies**[3]** and their difficulties are talked about in detail. It is accepted that this paper would profit both cloud clients and scientists in defeating the difficulties confronted.

The location of the information relative to the available computing resource is one of the main challenges for effective software scaling. In terms of performance, co-locating data and computing is ideal, particularly for data-intensive applications. Of various reasons, however, this is not always feasible. For example, data may be placed with minimal or overloaded computing resources in the processing nodes of the local organizational infrastructure of the user (e.g., a cluster or desktop grid). The user facing cut-off time requirements may also want to use a public cloud provider's on-demand computing resources**[4]** to can the application's execution time. In the above specific scenario, it may not be desirable for the user to move the entire dataset to the cloud as the data transfer time the rule over the performance benefit resulting from using external CPUs due to data size and system data transmission.

A video Transcoding platform can provide video Transcoding resources with heterogeneous QoS requirements dynamically. We adopted the pre-emptive resume priority discipline to design a multi-priority Transcoding mechanism that can improve resource use without affecting the delay sensitive QoS. Consider multi-core / multi-mode video transcoding scheduling and hardware-accelerated video transcoding resource provisioning **[5]**.

Cloud computing propelled model conveys data innovation benefits in which resources are recovered from the web through online devices and applications. The significant difficulties in distributed computing are privacy, resource allotment and provisioning of resources. In cloud computing, mechanism for provisioning resources is necessary to provide cloud buyers a lot of registering resource for handling the occupations and putting away the information. Cloud providers can provide two resource service plans to cloud shoppers, in particular present moment on-request and long haul reservation plans. Proficient resource provision which can ensure the acceptable distributed computing administrations to the end client, establishes the framework for the accomplishment of business rivalry. The significant point in resource provisioning is greatest execution in least time and diminishes the measure of

information move with least cost. At the point when the remaining task at hand of administrations increments quickly, Quality of Service (QoS)[6] can corrupt the facilitated request and the Service Level Objective (SLO) may be violated.

The issue of task scheduling on a disseminated domain has grabbed the attention of analysts. Task scheduling **[7]** is viewed as a fundamental issue in the cloud storage environment by considering various variables such as consummation time, complete expenditure on the execution of each user's undertaking, resource utilization, control utilization, and non-critical failure adaptation. A bi-target enhancement challenge is the issue of finding the right trade-off between the pacing of the targets and the vitality consumed by a prioritized parallel process. The answer for this issue is a lot of Pareto focuses. Pareto arrangements are those for which one goal can only be improved by compounding another goal in any event. Along these lines, rather than an interesting answer for the issue, the answer for a bi-target issue is a (conceivably unbounded) arrangement of Pareto focuses. Errand planning has been demonstrated as a NP-complete problem.

Task scheduling **[8]**assumes a key job in cloud computing frameworks. Undertaking scheduling on the basis of certain requirements is impossible under a lot of standards and guidelines that we can infer as an arrangement between cloud users and providers. This agreement is only the nature of administration that the client needs from the vendors. Giving great nature of administrations to the clients as per the understanding is an unequivocal assignment for the suppliers as simultaneously there are countless errands running at the supplier's side. The scheduling problem of the undertaking plan can be seen as identifying or looking through an ideal mapping / task for subtasks of different assignments over usable resource arrangements (processors / PC machines) so that we can achieve the ideal goals for errands. By performing relative investigation of the various algorithms for their reasonableness, attainability, flexibility with regards to cloud situation, after that we attempt to propose a hybrid solution that can be used to further improve the current stage. With the goal cloud providers can be allowed to give administration a better nature.

Cloud computing provides on-demand processing and capacity of storage with high performance. Several computing models have vowed to convey this utility computing. Cloud computing is one such dependable processing paragon. We check the viability of the proposed**[9]**multi-objective hybrid bacteria foraging algorithm (MHBFA) explicitly its job in solution decent variety and quality, consistency and dependability. Be that as it may, the rising vitality utilization of cloud data centres has turned into a noticeable issue. Task scheduling is a significant advance to improve the general execution of the cloud computing. Customary observing and the executives systems are intended for big business situations, particularly a brought together condition. Be that as it may, the enormous scale of heterogeneous resource provisioning in different data centres places real difficulties in managing and checking components.

3. HRPC PROPOSED SYSTEM

In proposed system the inputs are assumed as end users, in which they utilize the services like Amazon Shopping, Educational Research, Online Business, News Feed, Healthcare, etc.,. The tasks which are collected from the data centres are analysed and scheduled. AHP uses a structural technique to rank user alternatives consistently based on the hierarchy. Alternative ranking is done based on response time, task costs and reliability. The assumptions are made to represent user requirements. The weights assess the attribute's relative priority. The resources are allocated according to AHP calculated priorities. Using ACO, AHP reassigns priorities and the re-assignment considers the resources available as the best possible allocation. The simulation results articulate that, as compared to the existing algorithm, the proposed algorithm significantly minimizes network latency, execution cost and time, ensures the optimal use of resources.

In Fig 1, the users requests for a task(T1…Tn), Where the assigned tasks has separate Broker Id. The tasks are mapped into the data centres (DC1…DCn). The cloud providers will ensure the spaces in the resources and then allocate the tasks. The initialization is created for each data centre, VM, and broker. For allocating the task Ant Colony Optimization is used to choose the optimal path. The two steps carried out in Ant Colony Optimization,

i. Pheromone Update
ii. Next Hop Calculation

Analytical Hierarchical Process is to rank the process based on weights. In this, a pair-wise comparison matrix is generated from that the data inconsistences are removed and then the Consistency Index (CI) and Consistency Ratio (CR) are calculated. The tasks are allocated optimally and the response time and cost are reduced.

International Innovative Research Journal of Engineering and Technology
ISSN: 2456-1983 Vol: 5 Issue: 3 March 2020

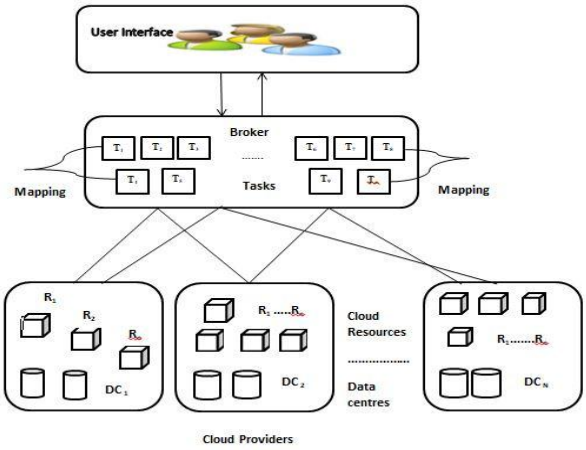

Figure 1. HRPC Architecture Diagram

4. EXPERIMENTAL RESULTS

The Cloud-Sim provides cloud environment, File size, Million Instruction per Second (MIPS), Bandwidth (BW), Random Access Memory (RAM), Virtual Machines (VMs) configuration, Host and Data centre, as well as cloud environment for physical components. Based on user requirements Cloud-Sim customizes scheduling in scheduling environments. This simulation tool is used to check the response time and cost. Two algorithms was implemented using Cloud-Sim simulation. For experiments that include VMs, data centres are used. We used up to six different virtual machines (VMs) and different cloudlet numbers (tasks).The performance is evaluated based on these two algorithms for scheduling cloudets, depending on the cost and response time. The datasets are taken from kaggle for medical diagnosis.

Table 1. Experimental Results

S.No	Performance Factor	Existing System	Proposed System
1	Time	The execution time was too high	Execution time is reduced by considering optimal path.
2	Cost	Here the cost was not optimal	Reduction of cost is done
3	Multi –Tier Application	Automatic bottleneck identification and resolution hosted in a cloud multitier web application.	Automatic Identification and resolution of bottlenecks hosted in a cloud multitier web application

Execution Time

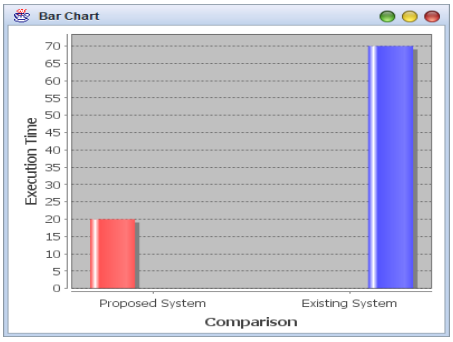

Figure 2. Execution Time

Cost Analysis

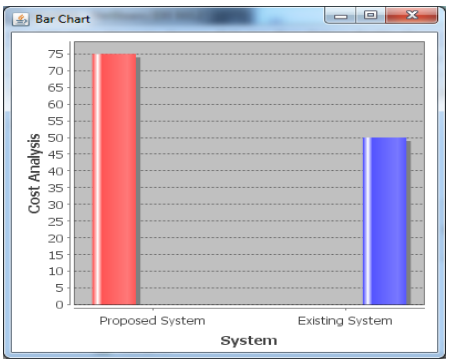

Figure 3. Cost Analysis

5. CONCLUSION AND FUTURE ENHANCEMENT

Cloud computing is a distributed based computer paradigm which is used by users to get good quality with less cost. As task scheduling is a challenge in cloud computing, different algorithms have been suggested and applied to get better results regarding utilization of system resources, response time and satisfaction of user demands. Thus our proposed optimal resource Provisioning provides better efficiency and cost for each task. Therefore, it reduces the response of time of Virtual Machines (VMs). We have analysed workload of each jobs. The primary goal of providing services is to achieve cost savings from the cloud service provider's perspective and from the cloud user's perspective. Results show that without increasing additional overheads our algorithm stands well.In the future, methods need to be introduced to make cloud services more effective in order to meet QoS and reduce SLA violation as resources are dynamically provisioned hybrid clouds. Hence, such provisioning methods must also beused for both SaaS and IaaS users.

REFERENCES

[1] S. Chaisiri, B. Lee and D. Niyato, "Optimization of Resource Provisioning Cost in Cloud Computing", IEEE Transactions on Services Computing, vol. 5, no. 2, pp. 164-177, 2012.

[2] Pradeep Krishnadoss, Prem Jacob2, "OCSA: Task Scheduling Algorithm in Cloud Computing Environment", International Journal of Intelligent Engineering and Systems, Vol.11, No.3, 2018.

[3] Adel Nadjaran Toosi, Richard O. Sinnott, Rajkumar Buyya, "Resource provisioning for data-intensive applications with deadline constraints on hybrid clouds using Aneka", pp.765-775, 2018.

[4] W. Deng, F. Liu, H. Jin, and C. Wu." Smartdpss: Cost-minimizing multi-source power supply for data centers with arbitrary demand". International Conference on Distributed Computing Systems, 2013.

[5] Guanyu Gao, Yonggang Wen, Cedric Westphal, "Dynamic Priority-based Resource Provisioning for Video Transcoding with Heterogeneous QoS", IEEE Transactions on Circuits and Systems for Video Technology, pp. 1051-8215, 2018.

[6] Pradeep Kumar V, Kolla Bhanu Prakash, "QoS Aware Resource Provisioning in Federated Cloud and Analyzing Maximum Resource Utilization in Agent Based Model", International Journal of Innovative Technology and Exploring Engineering, Vol.8, 2019.

[7] I. Goiri, K. Le, M. E. Haque, R. Beauchea, T. D. Nguyen, J. Guitart,J. Torres, R. Bianchini. "Greenslot. scheduling energy consumptionin green datacenters". International Conference for High Performance Computing, Networking, Storage and Analysis, 2011.

[8] Jiangtao Zhang, HejiaoHuang, XuanWang. "Resource provision algorithms in cloud computing: A survey". ELSEVIER, pp. 23-42, 2016.

[9] Sobhanayak Srichandan, Turuk Ashok Kumar, Sahoo Bibhudatta, "Task scheduling for cloud computing using multi-objective hybrid bacteria foraging algorithm", Future Computing and Informatics Journal 3, pp. 210-230, 2018.

[10] Ankita Jain, Arun Kumar Yadav, Brijesh Kumar Chaurasia "A Proactive Approach for Resource Provisioning in Cloud Computing" , International Journal of Recent Technology and Engineering, Vol. 7, 2019.

[11] Babur Hayat Malik, Talia Anwar, "Investigating Methods of Resource Provisioning Mechanisms in Cloud", International Journal of Advanced Computer Science and Application, Vol. 9, 2018.

AUTHOR BIOGRAPHY

Dr. M.S.Anbarasi is currently working as Assistant Professor in the Department of Information Technology at Pondicherry Engineering College. Her area of interest mainly focuses Data Mining, Databases and Cloud Computing.

V. Jayasri is pursuing M.Tech degree in Department of Information Technology, Pondicherry Engineering College, and Pondicherry.