International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

Design of Enhanced Dynamic Resource Allocation Framework for Heterogeneous Cloud Environment

L. Jayasimman, V.Geetha Dhanalakshmi

Abstract*: Cloud is a heterogeneous environment which is used to provide computing infrastructure as a service to the users. Cloud services are delivered to the users based on the demand received from the cloud service users. The best quality of a business model is to satisfy the product demand on time. Cloud is a business model to enable computing infrastructure everywhere. The cloud resources are scheduled and allocated to the users based on the requirement. Resource allocation in cloud computing environment is suffered by lack of improper monitoring, measurement and on time delivery of resources. This may affect the business growth of cloud service providers. This paper proposes an enhanced dynamic resource allocation framework to allocate resources based on the demands as well as cost of the resources to the users. The demand is forecasted based on the previous request repository. Fastcaster analyzes previous demands on different time and predicts future demands. Resources are prepared and to be in a ready state for allocation. The paper also describes the procedure of the proposed dynamic resource allocation technique for efficient allocation of resources in the cloud. The main focus of this framework is to schedule and allocate the resources which are in demand by the users on time also considered cost of the resources. Framework initiates a new paradigm of a single point of service delivery for all cloud services.

***Keywords* : Resource allocation, Forecaster, Users Request, Demands, Dynamic Allocation**

I. INTRODUCTION

Cloud resources are virtually provided to the service demands from the users. Cloud is a place of high energy computing infrastructure which could be rented to the service users. Allocation of the resources is a vital task in the cloud environment [1]. Cloud is a business model which helps to small and medium scale enterprises to pick up their business through the cloud infrastructure. Cloud is attracted because of the characters like on-demand service and so on. On-demand service delivery is the most important quality of service for business promotion [2] [3].

Many cloud service providers are delivered their services based on a different subscription basis. Cloud service providers are the source of high-power computing, and each provider delivers different types of services.

Revised Manuscript Received on January 30, 2020.
***** Correspondence Author
L. Jayasimman, Associate Professor, PG & Research Department of Computer Science, Bishop Heber College, Trichy. Affiliated to Bharathidasan University,Trichy. E-Mail: simmanjaysee@gmail.com
V.Geetha Dhanalakshmi*, Research Scholar, PG & Research Department of Computer Science, Bishop Heber College, Trichy. Affiliated to Bharathidasan University, Trichy. E-mail: geethadhanalakshmi @gmail.com

Users can have links with more than one service providers for their entire computing requirements. Services consumed by the users are charged based on the service used by the users on a pay per usage basis [4]. Cloud users should aware of each service provide regarding their cloud services. All the cloud services are independent, and users want to communicate with each service provider, then they must conduct each of them separately for accessing the resources. If users have accessed so many cloud services then, they must maintain and monitor the registration and login details of each service in the cloud.

In a cloud environment, several numbers of cloud service providers are in action to provide services. There is no Single Point of Service (SPS) for all services [5]. If the cloud may have a single point of services, then it is more comfortable for all users to use and contact a single point to get all their required computing resources [6]. Cloud has much magnetism, but it lacks in many sides of its benefits. Cloud has a vast computing repository and lacs many users are requesting resources every time in a day [7]. Cloud lacks in allocating the resources in an efficient and optimal manner. Cloud resources should be utilized fully without any waste of resources. Allocating cloud resources based on the users' demand is a critical task in heterogeneous cloud computing [8]. This paper considered the challenge of allocating resources efficiently and have a enhanced dynamic resource allocation framework for providing cloud resource and provide services from a single point of service delivery in heterogeneous cloud computing [9]. The paper proposes a dynamic resource allocation framework with different entities. The availability of resources is calculated based on service delivery by the different cloud service providers. Resources from different cloud service provider are cumulatively maintained in the integrated heterogeneous cloud environment and demands for resources are predicted for scheduling the resources [10]. The paper also discusses an efficient resource allocation technique for the dynamic nature of the cloud environment. Resources are allocated based on the demands and cost of the resources.

II. RELATED WORK

Cloud computing is researched in a different direction to improve the efficiency of cloud usage and adoption. Researchers come with a different solution for resource allocation, but they also insist on developing an efficient method to improve resource allocation. This section discusses some of the related research work carried out by the researcher in cloud resource allocation. Rengasamy et al. [17] proposed a predictive based a resource allocation with enhanced load balancing.

This approach allows cloud service providers to be in ready state to allocate resource for eny events. The resource allocation is based on Cicada predictions and it uses C-Rule algorithm.

Resource allocation and energy efficiency of the cloud environment si also consider in the framework. Cicada is a predictor which is used to predict the workloads early to know the demands of the resources with the help of C-Rule algorithm. The framework is simulated in the cloudsim with the a rule based algorithm. The advantages of the C-Rule algorithm is to avoid the overloads. It is better to indentify overloads and C- Rule in cloudsim, it helps to find the requested resources for all requests. Cloudsim takes less computational energy for resource allocation and C-Rule algorithm makes allocation with minimal resources.

Jayaraman et al. [11] proposed high utilization strategies to increase occupancy and improve system utilization. The paper proposed a scheduler called DRIVE meta scheduler to evaluate workload to a different scenario, throughput, and short job duration workloads. It is developed to assign workload in distributed and federated computing environments. An economic market strategy is used in DRIVE to an abstract allocation of resources. This scenario reduces the necessity for dedicated infrastructure and distributes management functionality across participants. Apart from the occupancy and utilization developments the proposed strategies also have advantages for differing economic conditions — the substitute service providers showed more price uncertain than other strategies because of the decreased allocation ratio at the time of bidding strategy.

KC Gouda et al., [12] proposed an approach that allocates resource with minimum wastage and provides a maximum profit. The established resource allocation algorithm is based on different parameters like time, cost, No. of processor demand, etc. The established priority algorithm is applied for improved resource allocation of a task in the cloud situation used for the simulation of several structure or task in an effectual method. At the completion of resource allocation for different tasks then an illustration is given based on the performance evaluation of cloud computing. Cloud computing servers are running for allocating resources for the task and job submitted by the users. The cloud administrator plays a crucial role in efficient resource allocation because he decides the priority among the different user request.

Manish Pandey et al., [13] proposed a cost-based resource allocation approach which migrates the virtual machine from one physical host (physical machine) to another machine to evade the work overloaded and under load situation of the host. A cost-based resource allocation strategy will allocate the VM first which paid more to the provider for using the cloud resources. Hence, the cost metric is used to place the VM. First, they calculate the bill which is paid by the customer for using the VM and arrange all VM in descending order according to the cost. During the assignment of a VM to PM, they maintained a list which satisfies the VM requirements. Then they assigned VM to all available PM one by one and calculate the utilization of PM before and after assignment and maintain a list. Arrange all PM into avail PMList in ascending order according to their diff value and assign the VM to the first PM. Continue the process until all the VM placed. This proposed approach is applied in the cloud simulator tool CloudSim to assess the quality of the proposed approach performance.

Wang Yan et al., [14] discussed a model, which is the adaptive and completion time aware model (ACTA) which gives the job adaptive and completion time aware for resource allocation. Moreover, a strategy is proposed based on the model. Non-cooperative Nash equilibrium theory is used to build a strategy with the goal of aiming to optimize cloud resources allocation. Accordingly, an allocation algorithm based on the gradient projection is presented. With experimental studies, the algorithm for resource scheduling is demonstrated to improve the efficiency of jobs and the cloud computing system resource utilization when compared to the fair scheduling, random scheduling, earliest deadline-time first scheduling. Wei Ming et al., [15] proposed a resource allocation method based on ant colony algorithm, which collect the information of the resource, such as the consumption of each server in the data center, and the power of CPU processing, as well as the initial size of the Memory that not been placed cloud. The information, such as saving Node Id, processing power, Memory, and the ability consumption, are all maintained in a table. Assume that all of the capacity loss, which is provided for all the nodes by those left components of a single node, are the same. And assume that all the nodes have the same network connection, as well as the same storage interfaces. After submitting a service request to the network Master, the system will normalize matching parametric.

Mayanka Katyal et al., [16] proposed a scheme for resource allocation for workflow based dependent tasks. In this scheme, prior to the allocation of tasks, the relationship among tasks in a workflow must be modeled. Directed Acyclic Graph (DAG) is used to model dependencies among tasks in a given workflow. Once the workflow is modeled, it is essential to ensure that tasks in a workflow are scheduled on appropriate VMs. In the absence of an appropriate allocation method, the tasks that require the same data set for completion might be allocated to VMs on a different host. This will result in multiple transmission of the same data set to a different host. This, in turn, increases the volume of data transfer which consequently leads to large bandwidth requirement and increases the cost of usage — also the response time, delays due to massive data transfer.

III. PROBLEM DEFINITION

Cloud resource monitoring and maintenance is a vital task. Allocating resources to the user based on the demand and cost is the most important area to be addressed in the cloud. Cloud is dynamic in nature and its demand highly vary on time to time. It is a crucial challenge to allocate resource to dynamic demands with different cost. Demands are changed on time, but cloud resources are always ready to allocate but

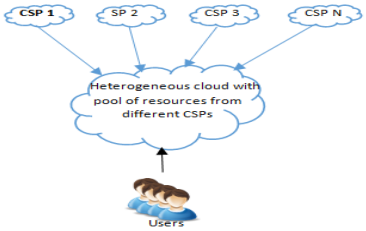

Fig. 1 Conceptual Diagram of Proposed Scheme

without knowing the demands of the resource, the allocation is not efficient, its lead to wastage of resources.

To avoid the wastage of resource allocation and to satisfy dynamic demands on time, it is necessary to develop a dynamic resource allocation framework to address the challenges in resource allocation in the cloud environment.

IV. METHODOLOGY

Cloud is constituted by a huge amount of resources and it is vary based on the cloud service providers. Cloud service providers are people who have the cloud infrastructure to provide the resources to users. Each cloud service providers may have to provide specific services or resources to the users. No cloud service provider who can provide all the service requested by the users. The proposed framework enables the cloud environment to deliver a single point of delivery to the users. Figure 1 shows the conceptual diagram of the proposed scheme. Users no need to approach different cloud service providers for each service. The proposed scheme provides an Integrated Heterogeneous Cloud Services Center (IHCSC) which provides all cloud services request by the users. Different cloud service providers are connected with the IHCSC and enable the users to get their service instantly. The proposed framework has different entities for monitoring and measuring the requested resources in IHCSC. Framework forecast the demands for allocating resources and check the availability of the resources based on the demand predicted from the requested repository in the framework.

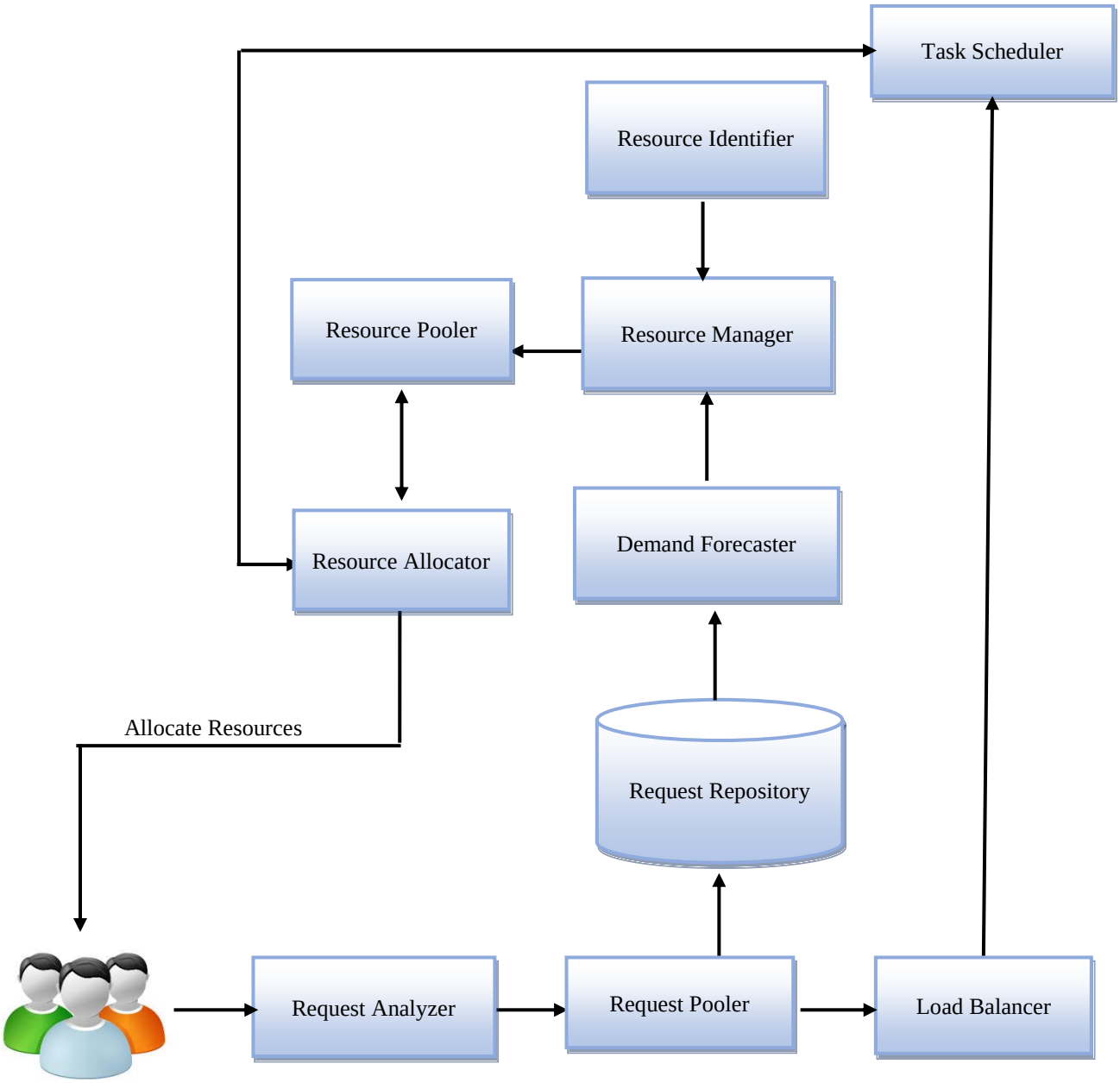

Fig. 2 Dynamic Resource Allocation Framework

V. RESOURCE ALLOCATION FRAMEWORK

The proposed framework has different entities which all support to schedule and allocate the cloud resources on time. The components in the frameworks are such as Users, Request Analyzer, Request pooler, Load balancer, Demand forecaster, Request manager, Request repository, Resource manager, Resource identifier, Task scheduler, Resource pooler and Resource allocator. Users requests are collected and analyst for identifying the demand for the resources. Based on the request analyst result, it can be defined that the highest demanded resources. Demand may vary on time. The proposed system gives maximum priority to the highest demanded resource and these resources are given priority to schedule for allocation.

Each entity is the proposed system is used for a specific purpose. Figure 2 shows enhanced dynamic resource allocation framework for efficient allocation of resources in heterogeneous cloud environment. The demand and cost are two parameters should be considered at time of allocating resource in the proposed framework. Demand denotes the highest priority of the resource requested by users and cost is used to select the resources from the resource providers. If a resource provided by more than one provider with different cost, then, it is analyzed for cost based selection.

Users initiate requests based on their requirements. Requests are may accessed different services like software, platform and infrastructure.

Request analyzer analyzes the different request given by the users. It categorizes the request based on the type and nature of the request.

Request pooler is a queue of request arranged for schedule and allocation. Request pooler is updated frequently based on the request generated by the users.

Request repository contains all the request which are allocated previously. It maintains the details of the full report of request received from and allocated to the users.

Demand forecaster is used to forecast the highest demanded resource. It is periodically updated with recent demand forecast details.

Resource manager manages the resources in the cloud data center. It receives a detailed report from resource identifier regarding the status of the available resource in each cloud service providers. The resource manager is responsible for adding and removing resources based on the request received and demands.

Resource pooler contains a pool of available resources which are scheduled for allocation. Scheduling is done based on the forecasting of demand by the demand forecaster. Task scheduler schedules the resources based on the availability of resource from the cloud service provider to the users' request.

Resource allocator allocates the resources based on the higher prioritized resources. Resources are on ready state based on the demand and cost.

The proposed framework initiates the Integrated Heterogeneous Cloud Service Center which has connected with different cloud service providers. It maintains a service catalog which shows the available resources from the different service providers. Figure 3 shows the connected cloud service providers in the framework.

Resource availability in each cloud service provider is frequently updated by the Resource Identifier. Resource manager gets the current status of resource availability of resources from Resource Identifier. If there is a highest demand for specific resource, then the resource manager sends an instruction to make provisioning by adding more resources in the resource pool.

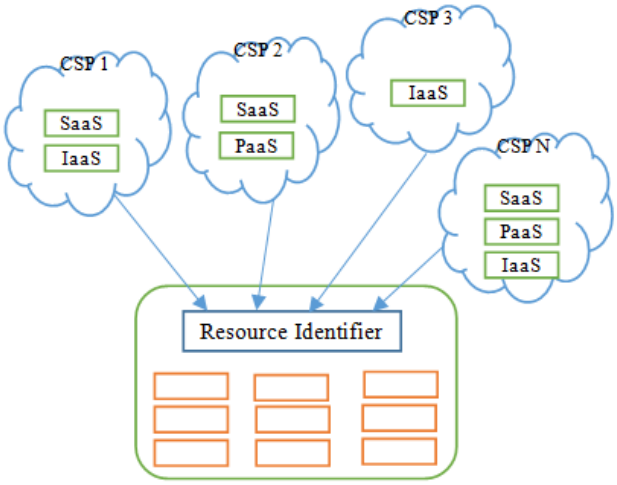

Fig. 3 Connected Cloud service providers with proposed framework

VI. RESOURCE ALLOCATION TECHNIQUE

The proposed framework concentrates on allocating the resources efficiently to users on time. The resource allocation technique is proposed in this section to consider the dynamic allocation of resources. The proposed resource allocation technique considers cost and priority-based resource allocation. The procedure involved in the proposed resource allocation is described in the below pseudo code. Initially, the total number of resource with all service providers are calculated. Find the total number of a resource request from the users. The highest demand resources are checked for availability. Resource available with the cloud service provider is allocated based on the cost and priority of the resources. Pseudo code for efficient resource allocation is given below.

Pseudo Code for Dynamic Resource Allocation (DRall) technique in the framework,

1. *Users request are analyst and find the total number of requests NRq.*

$$NRq = \sum_{i=1}^{n}(URq) \quad ...(1)$$

2. *Request are categorized for each resource to find the demand of each resource.*

$$DRSj=\sum_{i=1}^{n}(URqj) \quad ...(2)$$

3. *Resource availability is verified and find the total number of resources* NR_S.

$$NRs=\sum_{i=1}^{n} ARs) \quad ...(3)$$

4. *Find the highest demanded resource request from DRSj.*

$$HDr= \uparrow (DRSj) \quad ...(4)$$

5. *Find the cost C of each resource with each CSP.*
6. *Arrange the CSP list in an order based on minimum cost to the maximum cost of the resources.*

$$SCSPRs=MinCRs<MaxCRs \quad ...(5)$$

7. *Allocates highest priority resources based on the cost of each resource.*

$$RA=CRs\ [\uparrow (DRSj)] \quad ...(6)$$

8. *Check whether URqj<DRSj then,*
   Choose the resources with minimum cost and minimum resources, allocate the full resource from the CSP.
9. *If not, add more resources from the available CSP.*

The proposed resource allocation technique enables the cloud service providers to allocate the cloud infrastructure efficiently and make high utilization of cloud resources.

VII. SIMULATION SETUP AND RESULTS

The proposed resource allocation technique executes on the real-time cloud environment. It is simulated in window azure cloud setup. Three cloud services are hosted in Azure cloud platform, and these three services are connected with leading cloud service which maintains all entities in the heterogeneous cloud, and it maintains the availability of the resources from three cloud services. Each service hosted in the cloud is assigned by different cloud resources. Cloud users' request is received from the users, and it is analyzed highest demanded resources are allocated first. Demand is predicted based on the previous requests from the users. Three cloud servers with minimum configuration is rented for this simulation. All cloud services are coded in C# and hosted in the window azure platform as a service.

The proposed resource allocation DRall is measured for its efficiency by comparing it with existing resource allocation techniques. By using some sample data, the techniques are compared. Figure 4 represents response time comparison of existing and proposed techniques.

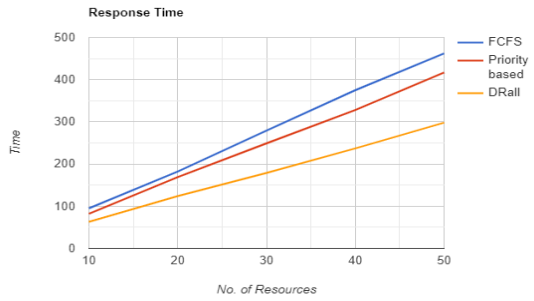

Fig. 4 Response Time

The framework provides efficient resource utilization compared with existing techniques. Figure 5 shows the percentage of resource utilization by proposed and existing techniques.

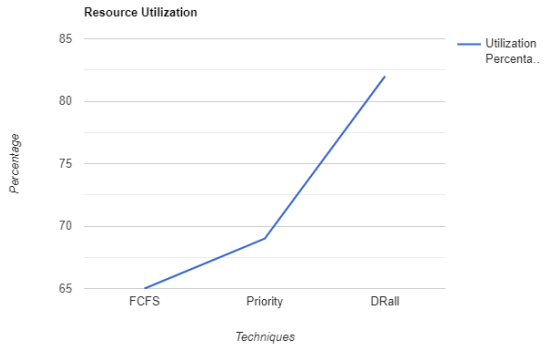

Fig. 5 Resource Utilization

VIII. CONCLUSION

The paper proposed a heterogeneous framework for allocating resources. Cloud has much attraction in providing service to the users. The providing services to users are a process of allocating the all requested resources with minimum waiting time. Allocating resources is a critical job in a cloud environment. The allocation may consider different parameters like time, cost, QoS and demand. This paper considered the demand and cost based resource allocation. All service providers are connected with the integrated heterogeneous cloud and availability of the resources are monitored and maintained by the resource identifier. Highest demanded resources given maximum priority for allocation and those resources are scheduled and provisioned for allocation. The resources selected based on the cost from different service providers. The resources are arranged based on the demand, highest demanded resources allocate first. The proposed system is simulated in real-time cloud setup using windows azure platform cloud. The proposed algorithm efficiently allocates the resources by considering the demands and cost. The results show that the proposed system is efficient in response time and resource utilization.

REFERENCE

1. Aleena Xavier T Rejimoan R., Survey on Various Resource Allocation Strategies in Cloud, 2016 International Conference on Circuit, Power and Computing Technologies, 2016, pp.1-4.
2. Luis M. Vaquero1, Luis Rodero-Merino1 , Juan Caceres1, Maik Lindner2, A Break in the Clouds: Towards a Cloud Definition, ACM SIGCOMM Computer Communication Review, Volume 39, Number 1, January 2009, 50-56.
3. Arockiam L, Monikandan S, Parthasarathy G. Cloud computing: a survey. International Journal of Internet Computing, Vol. 1, Issue 2, 2011, pp. 26-33.
4. Njakarison Menja Randriamasinoro, Kim Khoa Nguyen, Mohamed Cheriet, Optimized resource allocation in edge-cloud environment, 2018 Annual IEEE International Systems Conference (SysCon), 2018, pp. 1-8.
5. V. Geetha Dhanalakshmi and L. Jayasimman, Demand-Based Efficient Dynamic Resource Allocation Technique for Heterogeneous Cloud Environment, Journal of Applied Science and Computations, Volume VI, Issue II, February/2019, pp. 1290-1295.
6. L. Arockiam and A. Stanislas, A Unified Architecture for Optimal Resource Utilization in a Heterogeneous Cloud Environment, *International Journal of Machine Learning and Computing, Vol. 4, No. 4, August 2014, pp.383-388.*
7. Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, C´esar A. F. De Rose and Rajkumar Buyya, CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms, Software – Practice and Experience., Wiley Online Library, Vol 41, 2011, pp.23-50.
8. A. Stanislas, L. Arockiam,Match Making Service (MMS): A Centralized Architecture to Enhance Scalability in a Heterogeneous Cloud, International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 3, Issue 10, October 2013 ISSN: 2277 128X , 1128-1135.
9. Shweta T. Ukande, Prof. Rajesh C. Dharmik, Design and Analysis of Resource Allocation Techniques in Cloud Partitions, 2016 IEEE World Conference on Futuristic Trends in Research and Innovation for Social Welfare (WCFTR'16), 2016, pp.1-6.
10. Li Deng, ShengGang Bao, Fei Xu, Heng He, Yulin Ren, Chao Li, Resource Allocation Method Based on Combinatorial Double Auction Mechanism in Cloud Computing, IEEE Conference on Industrial Electronics and Applications (ICIEA), 2018, pp.801-806.
11. Jeyaram.G, Vidhya.V, Efficient Resource Allocation Strategies for Cloud Data Centers, International Journal of Advanced Research in Computer and Communication Engineering, Vol. 2, Issue 2, February 2013, pp.1230-1237.

12. KC Gouda, Radhika T V, Akshatha M , Priority-based resource allocation model for cloud computing, International Journal of Science, Engineering and Technology Research (IJSETR), Vol. 2, Issue 1, January 2013, pp. 215-219.
13. Manish Pandey, Sachin Kumar Verma, Cost-Based Resource Allocation Strategy for the Cloud Computing Environment, 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT) 2017, July 3 -5, 2017, pp. 1-7.
14. Wang Yan, Wang Jinkuan, Han Yinghua, and Wang Xin, Adaptive Resource Allocation Strategy in Cloud Computing Environment, The 9th International Conference on Modelling, Identification and Control (ICMIC 2017), Kunming, China, July 10-12, 2017, pp. 70-75.
15. Wei Ming, ZhangChunYan, Qiu-Feng, CuiYu, SuiQiangQain, DingWanBing, Resources allocation method on cloud computing , 2014 International Conference on Service Sciences, CPS, 2014 IEEE, 199-201.
16. MayankaKatyal, Atul Mishra, A Novel Resource Allocation Technique for Logically Dependent Tasks in a Cloud Computing Environment IEEE 2015 2nd International Conference on Computing for Sustainable Global Development, pp.1362-1366.
17. R.Rengasamy and M.Chidambaram, A Novel Predictive Resource Allocation Framework for Cloud Computing, IEEE International Conference on Advanced Computing & Communication Systems (ICACCS), 2019, pp, 118-122.
18. N. U. Ekici, K. W. Schmidt, A. Yazar and E. G. Schmidt, ACCLOUD-MAN - Power Efficient Resource Allocation for Heterogeneous Clouds, IEEE Signal Processing and Communications Applications Conference (SIU), Sivas, Turkey, 2019, pp. 1-4

AUTHORS PROFILE

Dr.L.Jayasimman, is working as Asst.Prof Dept. of Computer Science, Bishop Heber College (Autonomous) , Tiruchirappalli, Tamil Nadu, India. He has 16 years of experience in teaching and 5 years of experience in research. He has completed his MSc in St. Joseph's College, Tiruchirappalli and M.Tech in Bharathidasan University and also completed his PhD degree in Bharathidasan University, Tiruchirappalli in 2014. He has attended many International and National Conferences, Seminars and Workshops. He has published 60 research articles in the international/national conferences and journals. He had delivered more than 10 lectures in various National level seminars, Symposium and Conferences. His research interest is cloud computing, cloud security, IoT and web technology.

V.GeethaDhanalakshmi, is pursuing her Ph.D. in Computer Science from Bishop Heber College affiliated to Bharathidasan University, Tamil Nadu, India. She is currently working as Asst.Prof, Dept. of Computer Science, Srimad Andavan Arts and Science College (Autonomous) , Affiliated to Bharathidasan University, Tiruchirappalli, India. She has 12 years of teaching experience. She has published 3 papers in UGC approved journal. She has attended many seminars and conferences. Her research interests are cloud resource allocation and memory management.