Volume 8, No. 5, May-June 2017
ISSN No. 0976-5697
International Journal of Advanced Research in Computer Science
REVIEW ARTICLE
Available Online at www.ijarcs.info

# Time and Cost based Resource Provisioning Mechanism in Cloud Computing

Aishwarya Soni
Department of Computer Science and Engineering
MMMUT, Gorakhpur UP, India

Muzammil Hasan
Department of Computer Science and Engineering
MMMUT, Gorakhpur UP, India

***Abstract:*** Cloud Computing is novel technology which provides the computing resources to the users with on-demand access to the resources such as server, network domain and virtual machines. Computing resources offered to the user by Cloud Provider using Resource Provisioning Mechanism (RPM) in an optimal way of maximizing the benefits of all stake holders in cloud computing. RMP here refers to the techniques to provide the maximum utilization of resources in context to cloud provider and minimize the financial cost of services in context of user. So, it handles the issues and interest of both cloud provider and user. How to effectively implement RPM is a major concern in cloud computing environment. There are various types of PRM used now a day each having its own merits and demerits. Here in this paper we are trying to evaluate existing techniques on Cost based and Time based resource provisioning mechanism. We will study the pros and cons of all the popular mechanism proposed under these two categories and their performance on various Quality of Service (QoS).

***Keywords:*** Cloud Computing, Resource provisioning Mechanism(RPM), Quality of service(QoS), Availability, Service Level Agreement (SLA).

## I. INTRODUCTION

Cloud is defined as group of hardware, server, storage and network that facilitate the deliverance of computing service. Cloud services consist, provisioning of infrastructure, software and server via the Network based on customer requirements. NIST defined the Cloud Computing are as follow"cloud computing is the model that intended for enable suitable on-demand network access to a common pool of configurable computing resources (e.g. Network domains, servers, software, virtual machines and applications) which can quickly provided and free with nominal managing efforts or cloud service provider communication"[1].Cloud computing is method of provisioningnumerous of computing services on virtual machines which stay over the huge physical machine pool that reside in them. It raised into center of attention merely when we think about that which computer resources is constantly required technique that raise capacity or include special capability in existing surroundingswith no need to invest in original infrastructure, prepare latest personnel or prepare the licensefor new software. But it offered an enhanced explanation.

It is one of the most potential technologies in IT industry in which the computing resources is moving into computers. This is grown to be one of the famous words in the computer trade. This foundation concept is, fairly simple, to facilitate the huge computing resources which we require will exist in everywhere in cluster of computers and as per need we would connect and utilize them properly. The service provider in cloud model still works on the infrastructure in its own facilities. This cloud services includes the infrastructure as a services (IaaS), platform as a service (PaaS), software as a service (SaaS), storage as a service (STaaS), security as a service (SECaaS), and many more [2]. Therefore, various services offer to the customer with resource provisioning techniques.

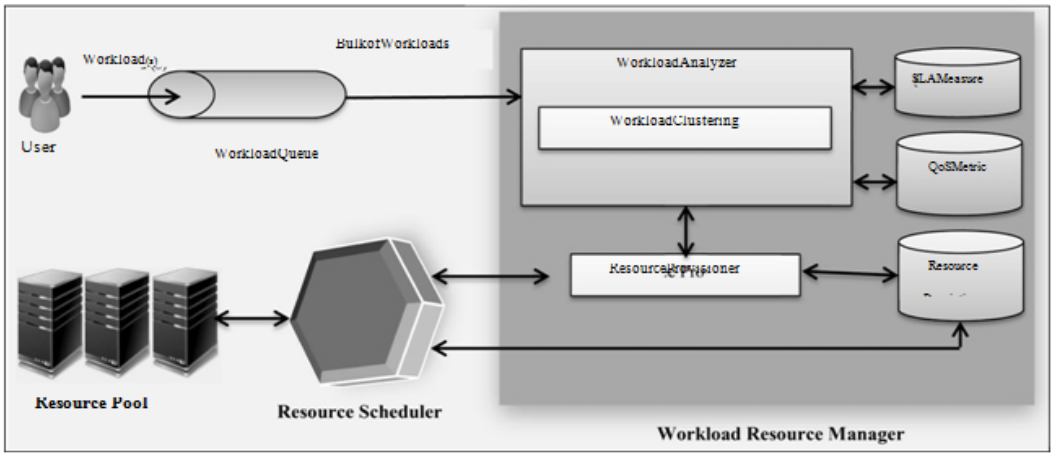


**Fig .1 Cloud Resource Provisioning [3]**

Resource management in cloud computing describes all the activity such as resource, provisioning, resources scheduling, resources and monitoring of resources. But all of them resources provisioning is challenging task because of lack of enough resources in cloud computing [3]. The appropriate resources provision to the workload need by

application depending on the quality of services (QoS) metrics in cloud computing. To provision the adequate resources to the workload with QoS metric is difficult task. Moreover the execution time minimization would consider in this task also a problem in existing research issue.

First of all, cloud users submit their request to broker for workload execution in the form of workload details based on QoS metrics. The broker matches the available resources to requested workload and decided how to provision the resources either based on QoS or not. After that scheduling perform by scheduler requested by broker. Additionally, the broker removes the extra available resource in pool of resources based on need of application. As in fig.1 Huge of workload submitted by users for execution and processed information stored in Workload queue. Workload analyzer(WA) given all the information of available resources based on QoS and SLA measures. Whenever the resources requested are submitted, Workloadanalyzer analyzes the available resources according to defined QoS constraint and characteristics. Then form a different cluster of prepare on the basis of CPU size, memory, cost, type of resources and number of resources [3,4]. Therefore,ResourceProvisioning means the selection, deployment, andrun-time management of software (e.g., databaseserver management systems, load balancers) andhardware resources (e.g., CPU, storage, andnetwork) for ensuring guaranteed performance forapplications. This is helpful to minimize the financial cost in aspects of users and maximize the resource utilization in perspective of resource provider. The resource provisioning used two allocation method in order to efficiently use of resources without violates the SLA and meet the Qos parameter. The other important factor or constraint in resource provisioning is Power consumption.In order to this constraint, it must reduce Power and power dissipation also on VM placement. There should be Technique to avoid access power consumption.

The desire goal in Resource Provisioning is to minimize the cost of rent in aspect of cloud user and efficiently utilization of resource in aspect of cloud providers. In order to allocation of resources there are two techniques Static Provisioning and Dynamic Provisioning. Either used the Static or Dynamic allocation of resources in perspective of requirement of resources. As per concern provider and users perspectives, requirements, outcomes and risks to compare the various resource provisioning techniques.Remaining sections are as follows: In section II, we have presented the, type of resources provisioning techniques along with the discussion of parameters in resource provisioning in cloud computing. In section III covers on available resource provisioning strategies.Section IV we have focused on comparison of resource provisioning method followed by section V the conclusion. This survey has not addressed any new resource allocation methodologies.

## II. RESOURCE PROVISIONING TYPE

They are classified on the basis of application needs which are as follows;

Static Provisioning.It is easy to configure in cloud computing to provision the computing resources. The computing resources needed by application or users have unchanged workloads. It does not vary and usually users request the required resources in advance before delivery of services. The users are charged on monthly basis.It is efficiently work and easy to prepare the resources in advance of start of services.

Dynamic Provisioning.Workload or computing resources needed by application is changed dynamically. It means requirement of resources is not fixed while it varies with requirements of application. The dynamic provision of resources is tough task compared to static provisioning.

Parameters for Resource Provisioning [5]

There are several parameters are as follows.

Response time: Minimum time require inexecuting the workload respond to whenever the users submit their request for workload.

Minimize Cost: As per users concerns the cost of computing resources they required must be in reasonable cost.

Revenue Maximization: As per cloud providers concern, aims to maximize their expected revenue in provisioning of resources.

Fault tolerant: Provision of resources must be fault tolerance refer able to continue to provision the services even any nodes get failure.

Reduced SLA Violation: Resources provisioning algorithm is designed in such a way that reduced the SLA violation.

Reduced Power Consumption: Power consumption required in migration technique and placement of Virtual machine would be lower.

## III.RESOURCE PROVISIONING MECHANISM.

Resource management contains all the activities such as resource scheduling, resource monitoring, and types of resource provisioning, resource provisioning mechanism. It describes all important characteristics in efficient resource utilization. Any kind of Resource Provisioning Mechanism required important characteristics like cost, time and energy. The provisioning of resourcedepend on the application requirement of QoS [16]. RPM play a key role in provision the appropriate workload to application. Based on the requirement of user's workload RPM adopts the dynamic actionswhich is more efficient than static RPM. It is designed in such a way to avoid underutilization and overutilization of resources [15]. There are various types of RPM, discuss the time based and cost based RPM. Most of the RPM widely used the dynamic or distributed features. Table 1 present the comparison on their common features of Cost Based RPM and Time Based RPM.

Cost based RPM

Several authors have been done research on cost feature of Resource provisioning mechanism. Abdullah M et al.[6] present the Divisible Load Theory (DLT) RPM is used to minimize the execution time ofuser applications,meet to satisfying the QoS constraints and maximize profit described by users while on executing the workload. This approach's reduced execution time and cost. Although the issues related to this approach is communication overhead and not able to handledynamic workloads. Hwang E et al. [7] investigate cost-effective provisioning of resource of Map-Reduce application with using deadline constraints, since the MapReduce programmingrepresentation is helpful and dominant in developing the data-incentive applications. It is used two cloud resourceprovisioning approach: firstly listed aspricing policies and secondly usedthe deadline-

aware taskspacking. These approaches minimize cost of Virtual Machines (VM) as well as meets deadline.But it is only adequate for the MapReduce applications. Additionally, the integration of RPM and workflow technology aredifficult task make thecomplex process to achieve the right requirement of the cloud resources which is essentialfor the execution of workflows with thelowest execution cost as well as highest resource utilization. Byun et al. [8] presented the framework which is based on workflow application that automatically executes the workloads and find the least requirement of resources to execute the application with in deadline. The mechanism in[8] able to reduce VMs cost, minimize the makespan and also satisfying deadline constraint but it is unable to deal with the dynamic workloads. Malawski M et al. [9] presented the research framework which integrates the dynamic and static methods. It is better deal with the resource provisioning as well as scheduling by executing the application with in their deadline and budget constraints.It reduces the provisioning delay and failures rates. But it is able to execute only similar type's workloads. Ming et al. [10] presented a framework that scales the resources automatically based on QoS and performance requirements of workloads and complete the execution within their desirable deadline. But it is not efficient mechanism for multi-tier applications. On the basis of abovementioned literatures, cost based RPM classify in to following categories as shown in fig 2.1. Multi QoS based Resources provisioning concern the different types of QoS parameters like time, energy etc. in a cost based RPM. While the virtualization based RPM is able to deal with efficient cost based provisioning of resources. In application based RPM, consider the three types of application like adaptive, data streaming and scientific workflow applications that is deployed for efficiently Provision the resources for cost effective. In Time based Application considers the execution time as secondary parameter of QoS. To improve resource utilization by avoiding the overutilization and underutilization, scalability based RPM is used for efficient provision of resources.

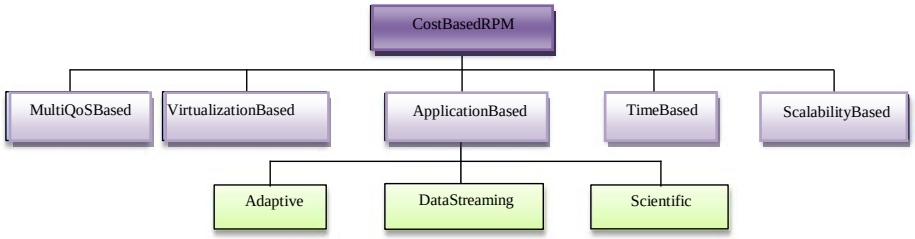


**Fig. 2.1 Cost based RPM Classification**

Time based RPM

Several Authors has been done their research work on time based RPM.Abrishami S et al. [11] presented the framework which is able to provisioning as well as scheduling large workflows named as IaaS Cloud Partial Critical Paths (IC-PCP) and IC-PCP by Deadline Distribution (IC-PCPD2) constraints. It reduced the entire computation time but it is unable to estimate accurate transmission and execution time. Buyya et al. [12] proposed a robust RPM which contain the resource allocation policies that executes the workflow in heterogeneous cloud resources. It is used to minimize the VMs cost and makespan time. Gao et al. [13] presented the framework of RPM which reduce the VMs cost of users application. It also improved the efficiency of energy and executes the workload within deadline constraints. There is no SLA violation and easily handle the workloads of multi end user environment but the admission controls is tough task. Moreover it reduces the power consumption and able to handles large scale workloads but it is not suitable for real time application.Vecchiola et al. [14] proposed the framework that concerns the QoS constraints and minimizes the makespan time by using resources from different service providers. It reduces the execution time but not allow the data intensive HPC application in presenting techniques.On the basis of abovementioned literature, time based RPM classified in to following categories as shown in fig 2.2. Time based RPM considered the three types of constraints such as deadline, budget and energy. Deadline based resource provisioning ensures that provision the resources to the user accordingly the urgently requirement of resources and characteristics of their submitted workloads. While in budget constraint, the budget specified by the users and resources are provision within deadline. To improve the efficiency of energy and utilization of resources, provision of resources to the users within deadline that meet the users' requirements.

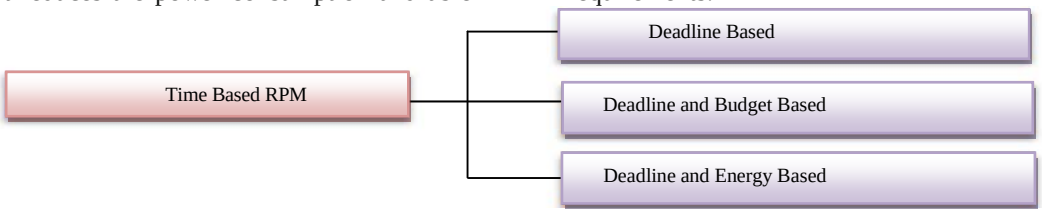


**Fig.2.2 Time based RPM Classifications**

## IV.COMPARISON OF RESOURCE PROVISIONING MECHANISM

Comparison of RPMs based on Cost based and Time based using the following the traits such as categories, scalability, merits, demerits etc to find the effective RPM.

**Table 1.Comparision of Cost based and Time based RPM**

| Resource Provisioning Mechanism | Category | Objective Function (goal) | Provisioning factors | Scalability | Merit | Demerit |
|---|---|---|---|---|---|---|
| Cost-based RPM | Multi-QoS based | To Help in reducing the processing cost. | Job deadline and delay cost | No | Minimize time and cost | Increases Communications overhead and unable to handle dynamic workloads |
| | Virtualization based | Reduced the cost of VMs. | Resource Utilization | No | Reducing the cost of VM and satisfy the deadline constraints. | Suitable for only Map Reduce Applications |
| | Application based | To Reduce total cost by find appropriate Task schedule. | Capacity of Resource and total cost. | Yes | Reduces the cost of resource and makespan time, Satisfydeadline constraints and Perform better. | Not able to handle Runtimeworkload |
| | Time based | Help to reduce time and cost. | Budget and deadline constraint. | No | Noprovisioning delay and lesser failure rate | Not able to handle heterogeneous workloads and does not concern transfer and data cost. |
| | Scalability based | Help to reduce users cost | Cost and deadline constraints | Yes | Reduced cost, no delay and satisfy deadline constraints. | It is not suitable for multi-tier applications |
| Time based RPM | Deadline based. | Reduced makespan of workflow | Deadline and normalized cost | No | Less Computation time | Does not estimate the accurate execution and transmission Time. |
| | Deadline and budget based. | To reduce Elapsed time. | Deadline errorand makespan time constraint | Yes | Improve the robustness andreduce the makespan time ofworkflow simultaneously | Increases the cost |
| | Deadline and energy based | To reduce energy and Execution time. | Deadline and energy cost constraint. | No | Easily handle the multi-user largescale workload | Admission controlis complex task |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Workflow based | To optimize resource Utilization. | Execution Time andCost. | No | Satisfy user defineddeadline andachievelower cost simultaneously. | Do not consider heterogeneous workflow Instances. |
| | Workload based | To reduce Computation time. | Time and cost. | Yes | Lesser cost schedule | Does not work in tight deadlines |

## V. CONCLUSIONS

When any user submits the workload request to the cloud provider for the resource allocation on cloud then Resource ProvisioningMechanism is used to identify the appropriate resources for submitted workload that ensures better performance, meet QoS constraints, reduces cost and time etc. It is used to improve response time, performance, save energy and QoS as well as adhere to SLA. The ultimate goal of RPM is to maximize profit from the both CloudService Provider's Perspective and CloudUser's Perspective to maximize benefit.RPM is one of the challenging tasks in cloud computing, there are several techniques of RPM, in this paper comparison are made on time based and cost based resource provisioning mechanism. Both have merits and demerits and different provisioning criteria and objective function. Cost based RPM has used different types of techniques in which some techniques helps to reduce the cost and execution time simultaneously.While the techniques used in Time based RPM has mainly focus onincreasing the utilization of resources and reduced the makespan time. It is also considered the energy constraints for better resource provisioning. Both of RPM has able to match the workload with available resource with their defined constraints. But the Cost based RPM has improved approach because of considering various objective that meet the users perspective aswell as providers perspective by meet the multi QoS constraints such as time, cost and energy.

## VI. REFERENCES


[1] Mell, P., &Grance, T. (2009). The NIST definition of cloud computing. National Institute of Standards and Technology, 53(6), 50.

[2] P.Mell and T.Grance, "The NIST Definition of CloudComputing", National Institute of Standards andTechnology, Information Technology Laboratory, TechnicalReport version 15 ,2009.

[3] Singh, S., & Chana, I. (2015). Q-aware: Quality of service based cloud resource provisioning. Computers & Electrical Engineering, 47, 138-160.

[4] Singh, Sukhpal, and Inderveer Chana. "QRSF: QoS-aware resource scheduling framework in cloud computing." The Journal of Supercomputing 71.1 (2015): 241-292.

[5] Nagesh, Bhavani B. "Resource Provisioning Techniques in Cloud Computing Environment-A Survey." IJRCCT 3.3 (2014): 395-401.

[6] Abdullah M, Othman M (2013) Cost-based multi-QoS job scheduling using divisible load theory inCloud computing. Proc Comput Sci 18:928–935

[7] Hwang E, Kim KH (2012) Minimizing cost of virtual machines for deadline-constrained mapreduceapplications in the Cloud.In:2012ACM/IEEE 13th international conference on grid computing (GRID).IEEE, pp 130–138

[8] Byun E-K, Kee Y-S, Kim J-S, Maeng S (2011) Cost optimized provisioning of elastic resources forapplication workflows. Future GenerComputSyst 27(8):1011–1026

[9] Malawski M, Juve G, Deelman E, Nabrzyski J (2012) Cost-and deadline-constrained provisioning forscientific workflow ensembles in iaas Clouds. In: Proceedings of the international conference on highperformance computing, networking, storage and analysis. IEEE Computer Society Press, p 22

[10] Mao M, Li J, Humphrey M (2010) Cloud auto-scaling with deadline and budget constraints. In: 201011th IEEE/ACM international conference on grid computing (GRID). IEEE, pp 41–48.

[11] Abrishami S, Naghibzadeh M, Epema DHJ (2013) Deadline-constrained workflow scheduling algorithmsfor infrastructure as a service clouds. Future GenerComputSyst 29(1):158–169

[12] Poola D, Garg SK, Buyya R, Yang Y, Ramamohanarao K (2014) Robust scheduling of scientific workflowswith deadline and budget constraints in Clouds. In: The 28th IEEE international conference onadvanced information networking and applications (AINA-2014), pp 1–8

[13] Gao Y, Wang Y, Gupta SK, Pedram M (2013) An energy and deadline aware resource provisioning,scheduling and optimization framework forCloud systems. In: Proceedings of the ninth IEEE/ACM/IFIPinternational conference on hardware/software codesign and system synthesis. IEEE Press, p 31

[14] Vecchiola C, Calheiros RN, Karunamoorthy D, Buyya R (2012) Deadline-driven provisioning ofresources for scientific applications in hybrid Clouds with Aneka. Future Gen ComputSyst 28(1):58–65

[15] Han, R., Ghanem, M. M., Guo, L., Guo, Y., & Osmond, M. (2014). Enabling cost-aware and adaptive elasticity of multi-tier cloud applications. Future Generation Computer Systems, 32, 82-98.

[16] Di, S., & Wang, C. L. (2013). Dynamic optimization of multiattribute resource allocation in self-organizing clouds. IEEE Transactions on parallel and distributed systems, 24(3), 464-478