

# Real Time Resource Allocation Methodology and Quality based Framework in Cloud Computing

Amit Kothari, PhD.  
Assist. Prof., Marwadi  
Education Foundation's Group  
of Institutions,  
"Shri Gopal Mandap Service",  
Nr. Nalanda School, Ranchhod  
Nagar Society, Rajkot

Nootan Padia  
Assist. Prof., Marwadi  
Education Foundation's Group  
of Institutions,  
"Shri Gopal Mandap Service",  
Nr. Nalanda School, Ranchhod  
Nagar Society, Rajkot

Maulik Parekh  
Assist. Prof., T. N. Rao College  
of IT,  
"Shri Gopal Mandap Service",  
Nr. Nalanda School, Ranchhod  
Nagar Society, Rajkot

## ABSTRACT

Cloud computing provides resources and services as per user's demand. User can pay for these services as per their use. Many cloud users can request no. of resources or services simultaneously. So resource allocation should be cost effective, reliable and efficient. There are lots of criteria which should be considered while allocating resources to the users. This paper shows resource demand categories and requirements to be considered while allotting resources to the customer. A quality based framework is proposed in this paper.

## General Terms

Real Time Resource Allocation in Cloud Computing and its Quality – based Framework.

## Keywords

Cloud computing, Cloud Deployment Models, Cloud Services, Virtual Machine, Resource Allocation, Resource Demand

## 1. INTRODUCTION

Cloud computing is a rent-based service where you can acquire network storage space and IT resources like software, hardware, servers, infrastructure etc. Cloud Computing is purely a set of shared computing resources and services delivered over the web. In CC, database and software are stored on the remote servers and applications can be accessed through web browser.

CC can be classified as private cloud, public cloud, community cloud and hybrid cloud. A public cloud can be accessed by any customer with an internet connection and access to the cloud space like Google Mail. A private cloud is recognized for a specific group or organization and limits access to just that group like Zimbra. A community cloud is shared among two or more organizations that have similar cloud requirements. For example, Airline management require same services. A hybrid cloud is a combination of more than one cloud i.e. combination of public, private or community cloud.

There are mainly two components of CC: front – end and back – end. Front – end component includes user interaction with client's network and application, and back – end component includes servers and data storage devices.

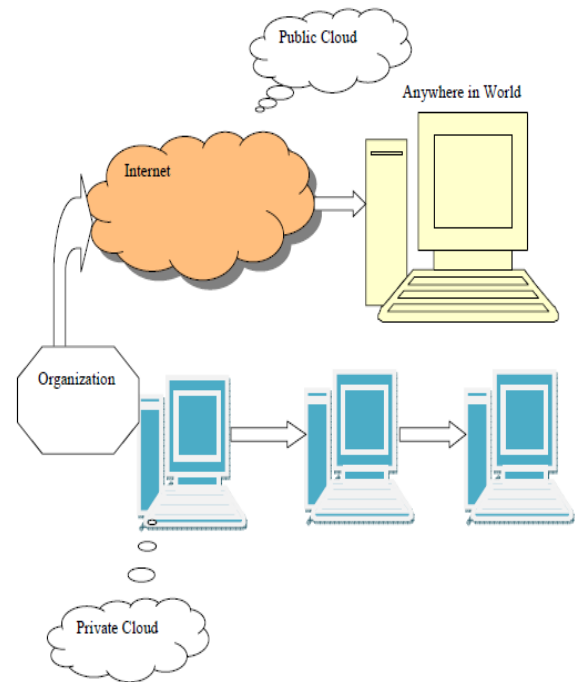


Fig1. Public and Private Cloud

Three types of services can be provided by cloud vendors – Infrastructure as a Service, Platform as a Service and Software as a Service. Analogous to P/I/SaaS, clouds may be hosted and engaged in different manners, depending on the use case scenario, corresponding to the business sculpt of the provider.

Table1. Cloud Services

Infrastructure- as – a – Service	Platform – as – a – Service	Software – as – a – Service
Physical Assets As Service	Resources on Demand	Services on demand
Lease the disk storage, n/w, power supply, and other resources	Used to design, create, run, and test the application	Built software are provided.

In resource allocation, suppose user logs in and demands for A, B and C resources, then cloud owner have ample no. of options to allocate these resources. Following are the requirements:

- a. Hardware requirements like CPU or RAM.
- b. Storage requirements like persistent storage or double buffering.
- c. Software requirements like operating system or computer language.
- d. Security requirements like free SSL or Authentication of user.
- e. Cheapest in cost, low response time and high performance is required.
- f. Energy efficiency

At the time of allocation of resources not only focus on resource demand also pays attention to cancellation of demand or service failure, because it costs a lot. Provider should provide resources when customers need and remove the resources when they don't need.

By considering these factors, cost – effective and fastest real time resource allocation process can be developed.

## **2. RELATED WORK ON RESOURCE DEMAND AND REAL TIME RESOURCE ALLOCATION**

Management of resources is the tough job in cloud computing. Resources may be CPU, memory, and disk capacity; bandwidth and latency (in server selection and network virtualization); worker cycles (in job scheduling) etc. Cloud vendors provide these resources as services to the customers. As no. of resource demand increases, the resource provision with high efficiency becomes problematic. Resource demand can be categorized as: a single user can demand single or multiple resources and multiple users can demand single or multiple resources. It will be easy task for resource manager when a single user demands for resources. Resource manager have to check the best available resources with low cost.

Problem is raised when multiple users demand single or multiple resources. At that time, resource manager has to allocate the resources on priority based. The same resource can be shared between multiple users simultaneously with the concept of virtualization. In real – time resource allocation, no. of users and no. of resource demands are not statistically defined.

Dynamically allocating/deallocating different compute resources from a cloud in its scheduling and during job execution is first introduced by Nephele data processing framework. [1] It also ensures security i.e. data can be accessed by its owner only. In this, two components are introduced: Job Manager and Cloud Controllers. In this, job manager schedules the job and cloud controller is work as an interface.

The dynamic resource allocation method for parallel data processing using novel turnaround time utility scheduling approach handles both high priority and low priority task. [2]

The semantically-enhanced resource allocator (SERA) framework is used for allocating resources from different

cloud providers using the semantic and multi-agent technologies. It is also defined how to dynamically distribute resources among virtual environments based on SLAs.[3]

When dynamic resource allocation is done with virtualization concept, at that time to avoid too much load, virtual machine should be migrated to other physical machines without any interruption. [4] Virtual machines are vibrant and dynamic. They can rapidly be reverted to preceding instances, stopped and resumed, moderately easily. They can also be cloned and flawlessly stimulated between physical servers. This dynamic character and latent for virtual machine spread out makes it difficult to accomplish and preserve unswerving security. Susceptible construction errors may be naively broadcasted. Also, it is hard to preserve an auditable trace of the security state of a virtual machine at any instance. In cloud computing environments, it is very much necessary to prove the security state of any system, irrespective of its location or immediacy to other, potentially timid virtual machines.

## **3. PROBLEM DEFINITION**

As per the study, most of the work is done either on job scheduling or response time. As no. of user increases, the demand of resources will also be increased.. At that time quality based resource allocation must be needed which fulfils all the requirements of customer. That should be energy saving, secure, and fastest. For e.g. if users want to download movie, then they have no. of options but they will select the option on the basis of security and speed. The same selection process should be done dynamically.

Sometimes it is also the case that, an enterprise has some resources on lease from one cloud provider. Now it needs such resources which are not available to that cloud provider and it has to contact other cloud providers. It cause much time and cost. Thus, the resource availability is also major problem.

Following qualities are highlighted in framework:

- a. High Data Transfer Speed
- b. High Storage capacity
- c. Distance
- d. Availability of Resources
- e. Security and Authenticity
- f. Low Power consumption
- g. Low Server Load

## **4. PROPOSED FRAMEWORK**

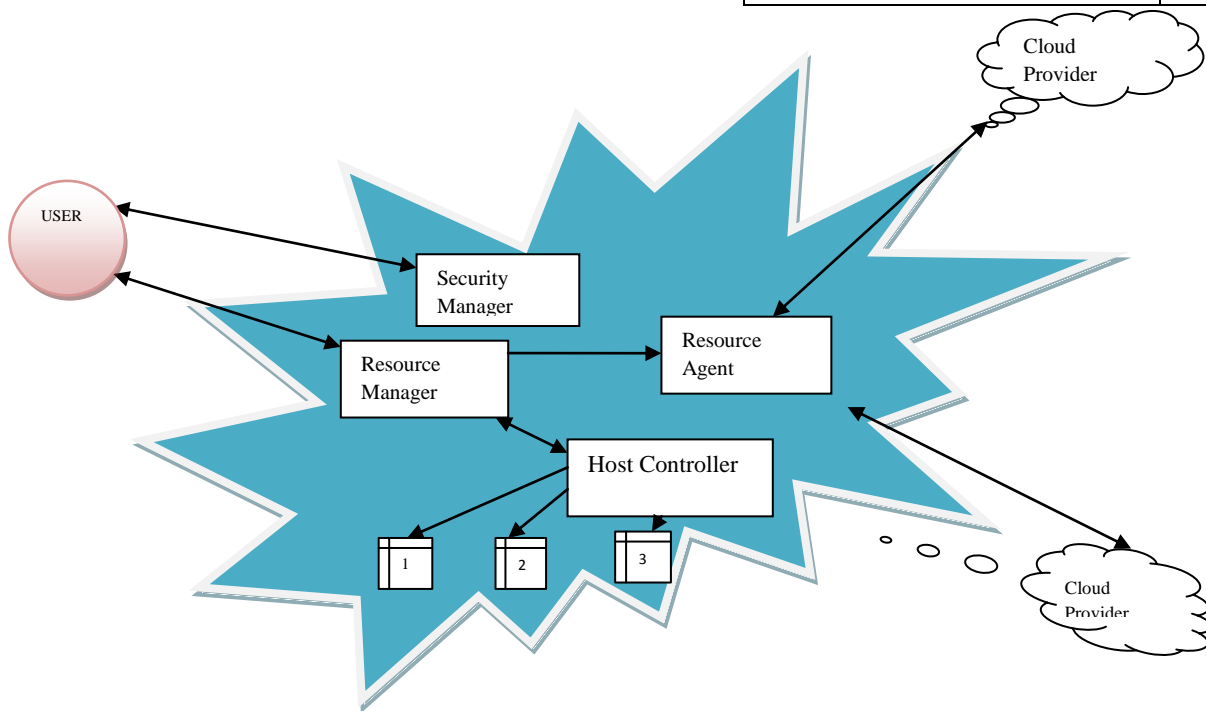
In this framework, five components are introduced: User Interface (UI), SM (Security Manager), RM (Resource Manager), RA (Resource Agent) and HC (Host Controller). User can login with UI and this login information will be confirmed by SM. SM also have firewall or SSL for security. After confirmation RM manage and provide the resources. Resource Manager also maintains the list of resources available in that cloud. If resources are not available in the particular cloud provider then Resource Agent will contact other cloud provider and get the resources. The Host Controller is a combination of virtual hosts. These virtual hosts will be at different locations and will be having different storage capacity, power and cooling map and data transfer rates. Virtual host is combination of various physical devices. For energy saving it should follow LLC.

Following are the steps:

1. User logins with username and password.
2. The security manager will check the user name and password and respond to User.
3. If matches, the user will request for resources.
4. The resource manager will do the following job :
  - a. Check the availability of resources in HC. If not available, then contact to RA.
  - b. If resources are available, then find the nearest hosts and select the best host from them. This host must be selected based on its capacity and speed.
  - c. Allocate the resources to the particular user.

**Table2. Framework Components Tasks**

Tasks	Components
User Interaction	UI
Managing Virtual Host <ol style="list-style-type: none"> <li>a. High Data Transfer Speed</li> <li>b. High Storage capacity</li> <li>c. Distance</li> <li>d. Low Server Load</li> <li>e. Low Power Consumption</li> </ol>	HC
Security and Authenticity	SM
Checking Availability and Selection of host	RM
Contact to other clouds to make resource always possible	RA



**Fig2. Framework**

## 5. CONCLUSION

After the implementation of proposed strategy, user will not have to worry about the scarcity of resource availability or problems like host failure. The cloud administrator also gets the same benefit. Resource manager can take the advantages of getting help from remote cloud using the “quality based methodology”. All the tasks will be done dynamically. User will get full transparency regarding the accessibility of resources without worrying about which server or cloud resources are provided. The purpose of the study is that the user will have their resources in a fastest and most secure manner with lesser power consumption.

## 6. FUTURE SCOPE

Every cloud is required to maintain the resource details of other available clouds and also needs to update the information periodically. Free resources should be available and provided on leased base. RM and RA can play an important role for such activity also.

In future, it will be decided that how to share the information between cloud through RA. Sharing will be periodically or on demand basis. RM will have to maintain the full status of available resources or only free resources/occupied resources whichever is required to send less information, and need to

have the strategy for maintaining the information of getting resources from other clouds on demand based.

## **7. REFERENCES**

- [1] Daniel Warneke and Odej Kao, "Exploiting Dynamic Resource Allocation for Efficient Parallel Data Processing in the Cloud", *IEEE transactions on parallel and distributed systems*, january 2011.
- [2] Venkatesa Kumar V. and S. Palaniswami, "A Dynamic Resource Allocation Method for Parallel Data Processing in Cloud Computing" *Journal of computer science* ,2012 Science Publications , pp. 780-788, 2012.
- [3] Ejarque, J., "SLA-Driven Semantically-Enhanced Dynamic Resource Allocator for Virtualized Service Providers" *IEEE Conference.*, pp. 8-15, Dec. 2008.
- [4] Chao-Tung yang, Hsiang-Yao Cheng, and Luan-Lung Huang, "A Dynamic Resource Allocation Model for Virtual Machine Management on Cloud", *Grid and Distributed Computing Communications in Computer and Information Sciences*, Volume 261, 2011, pp 581-590.
- [5] J. G. Koomey, C. Belady, M. Patterson, A. Santos, and K.-D. Lange, "Assessing Trends over Time in Performance, Costs, and Energy Use for Servers," Intel and Microsoft, Tech. Rep., 2009.
- [6] Armbrust, M., Fox, A., Griffith, R. et al. Above the Clouds: A Berkeley View of Cloud Computing. UCB/EECS-2009-28, EECS Department, University of California, Berkeley, 2009.
- [7] Zhikui Wang, Yuan Chen, Daniel Gmach, Sharad Singhal and Brian J. Watson, Appraise: Application-Level Performance Management in Virtualized Server Environments (Infocom) Spector, A. Z. 1989. Achieving application requirements. In *Distributed Systems*, S. Mullender