# Classification of Resources and Resource Management Techniques in Cloud Computing

Dr. G. Rama Subba Reddy
Associate Professor, Dept. of CSE, Mother Theresa Institute of Engineering and Technology, Palamaner, India

**Abstract**

Cloud computing is considered as on-demand service in which the applications to the data centers will be like pay-per-use manner. The predominant concepts in cloud computing are load balancing, allocation of resources and workflow scheduling. These are aimed at providing the cloud resources to cloud users in order to access in an easy and convenient way. Load balancing focuses to ignore the overload issues and nodes underload. The resource allocation focuses to enhance the performance levels by proper allocation of resources to use. The aim of workflow scheduling is to reduce the consumption of energy and cost and so on. Among these techniques, the key task related cloud computing environment is Resource management. Cloud computing enables cloud users to allow accessing the computing resources (comprising servers, networks, services and so on) with ease and less cost. It is required a flexible resource allocation strategy so that to assign those resources in proper way and also to meet the user's requirements. With increase in user requirement, the procedure of resource allocation becomes complex and high challenging one. Hence in this paper we concentrated on some different kinds of algorithms related to load balancing, resource allocation and workflow scheduling. We also presented the classification of resources in virtualized and non-virtualized environment.

**Keywords – cloud computing, components, virtualized and non-virtualized resources, resource management.**

## 1. Introduction

One of the recent technologies in todays is Cloud Computing, in which the clients are provided the services without considering time and location. The services will be regarding servers, storage and the platforms for the development of software. The software when connected with internet is considered as a "Cloud". Cloud computing is related with virtualization. The cloud computing comprises of the some deployment models and those are Private cloud, public cloud and hybrid cloud. An enterprise in private cloud model utilizes a model which is self-controlled model and runs inside its self-data centre. A 3rd party provider in public cloud enables computing resources to public through internet. In hybrid cloud model the collection of on-premises, private cloud and third-party public cloud services along with the dual platforms arrangements. The services provided by the Cloud computing are SaaS, PaaS and IaaS. SaaS is abbreviated as Software as a Service, in which a 3rd party provider allows the application and also provides the availability of those to the users via internet. In PaaS (Platform as a Service) service the 3rd party provider allows an application development form and requirements on its self-organizational framework and keeps it available to users by internet. In IaaS (Infrastructure as a Service) model the 3rd party provider will hosts the server and storage. Besides it makes them to be available to users via internet [1].

A Resource management includes with the allocation of resources, balancing of loads and also scheduling of workflows which are considered as major issues in the context of cloud computing. The load in the load balancing gets balanced in such a way that each node will gets loaded and it ignores the issues with nodes that are under loaded and over loaded. The strategy of resource allocation assigns the optimal resources to customers based on necessity and it leads in the improvement of QoS. To schedule the workflows it is needed to choose appropriate resources for the execution of tasks which influences the time of execution and also the cost. Every method has its own significance in Resource Management. The norms which must be avoided during resource allocation are Confliction between resources, resource lacking, resource fragmentation, high usage, less usage. Load balancing make sure that [2] load will be provided equally over nodes. In increase the performance levels of a system. It meets the user requirements. It leads to quick response to the requests by the users. The workflow scheduling includes [3] with Resource allocation and Task Scheduling.In simpler terms, a cloud is nothing but a gigantic pool where the accessing is easier and in includes with the useful virtual resources. The cloud computing is a model which is service providing and besides it allows many types of services to users [4]. The requirements will alters dynamically. Therefore the task of cloud computing is make use of requested services to users of a cloud. But to the cloud providers it is complex to offer all the requested services on time due to finite resources unavailability. The perspective of cloud providers is that the resources of cloud must be assigned in right way. Hence this is an important issue to satisfy the cloud users with QoS requirements [5]. Conventional techniques of resource management are inadequate for the cloud computing since they are on the basis of virtualization technology with decentralized property. The cloud computing proposed some novel challenges considering the resource management because of heterogeneity in the hardware capacity, on-demand service model, and pay per use model and gives assurance in meeting QoS [6].

## 2. Components

The systems in cloud resource management comprise of the following fundamental components and are shown in Figure 1.

### A. SLA Management

The SLA (Service Level Agreement) module generates an agreement between a client and service provider in order to provide the requested services to client. That agreement includes the client's requirements of resources such as CPU, storage and other VM's configuration model. Besides it includes with QoS requirements such as response time, task finishing time, operational cost etc. This SLA module interacts with the admin controller before making an agreement. When the agreement is successful then the request of users will be sent to scheduler.

**B. Admission Controller**

The client's request via SLA module is gets validated on the basis of requested resources availability and the other conditions stated in SLA. If those requirements are not satisfied then service provider rejects the request.

**C.Pricing**

Pricing is made over the dynamic activities on the basis of resource utilization. The consumption of time and cost of resources are evaluated and is charged to users in the real time.

**D. Scheduler**

The scheduler supplies the requested activities on existing VM or on Physical machine (PM). In such situation, there might be a queue to specify the job priority. At a time a single job is processed over a VM. The task of scheduler is to confirm on which the job is supplied either from VM or PM. The scheduler might get these details from load balancer. Whenever the user finished submitting the jobs, the module of pricing is invoked to generate the bill regarding resources usage by user and the interaction is done via SLA module.

**E. Load Balancer**

This detects the VM or PM which are over utilized and underutilized. It workloads gets balanced between the VMs in order to handle the usage of resources. Hence the energy is saved in this manner by using the resources which are idle.

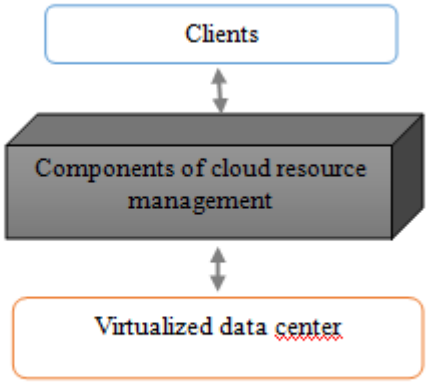


**Figure 1: Cloud resource management system**

## 3. Literature work

Under this section, we described the prevailing techniques in the cloud computing for resource allocation, load balancing and workflow scheduling.

### 3.1 Resource Allocation

The integration of the activities of a cloud provider is done to utilize and allocate many of the existing resources under a specific boundary of a cloud environment to achieve the requirements of a cloud application. The approaches for resource allocation are classified into different kinds by taking different parameters into sight. In paper[7] there recommended an algorithm OCRP along with a stochastic programming model applied decomposition paradigm. We considered multiple cloud service providers in this paper work. By applying this algorithm, the researcher concludes that costs get minimized during the allocation of resources. In paper [8] the author suggested Multi-objective optimization paradigm which is runtime friendly and also cost effective one. A technique known as Branch and Bound is applied to know the optimal situation of the resources. It can be done over the web applications under multiple scenarios. In paper [9] introduced an algorithm which is k-means to classify the tasks based in priorities. Here the priority is determined as if the job needs high amount of CPU and it will be executed first. In paper [10] there proposed a novel paradigm is Agent based Automated Service Composition which includes with the phases of processing the requests and composition of automated services. It optimizes the VM's cost when consumed by the on-demand services.

### 3.2 Load Balancing

In considering with the cloud computing, Load balancing is said to be one of the major issues. It minimizes consumption of time during service, response and waiting time of a job. In paper [11] introduced an algorithm Cuckoo Optimization for the load balancing. This involve in three stages. In the initial stage, it identifies the over utilized hosts. In the next stage it shifts from over utilized host to other one. In final stage, it choses VMs and also used for least migration policy. In paper [12], the author introduced an agent based paradigm to cloud platform. To balance the load, the load agent, channel agent and migration agent are used. Load agent controls the information policy. Channel agent and migration agent controls Transfer policy, selection policy and location policy. Migrate agents manages to mobilize to other data centers and also interacts to other load agent to obtain the status of current VM. In paper [13] there suggested dynamic load balancing with effective VMs configuration. It comprises of three tiers such as Web tier, schedule tier and Resource allocation tier. A web

tier accepts the requests from users. Schedule tier performs job scheduling on priority basis. Resource allocation tier enables the VM allocation. In paper [14] a load balancing paradigm is introduced in the mobile cloud computing on the basis of Max-Min Ant system.

### 3.3 Workflow Scheduling

The process to determine tasks and provide those tasks for the relevant resources to control the execution of work flow is known as Workflow scheduling. Besides it achieves the objectives regarding performance. In paper [17] Bandwidth Aware Task Scheduling (BATS) paradigm is introduced for the parallel task scheduling kind of applications. In paper [19] an algorithm for admission control and scheduling is introduced. It provides optimum resource usage and reduces the SLA violations. In paper [21] Bi-level advanced reservation approach where scheduling is done at local and global levels. The author applied some algorithms for the global scheduling and those are Critical Path Extraction and DAG partitioning algorithms. In paper [23] the extension to Max-min algorithm was proposed. It is valid for tasks allocation with an average execution time to reduce minimum finishing time of a resource.

### 3.4 Research Challenges

In the context of cloud systems, the research on RA is still at infant phase. Many prevailing issues are incompletely determined whereas the novel challenges continuing it emerging. Few of the challenging issues under this are described in the following:

**Migration of VM:** The issue of migration happens based on the user requirement to shift to other provider for getting better way to store the data.

**Control:** There might be a scarcity with control mechanism on resources since they are taken for rent by users from remote server.

**Energy Efficiency**: Because of large data centers emergence which has many computing operations are required for allocation of energy efficiency. Those centers results in the emission of carbon at high range.

**The Scheduling of Parallel Jobs:** In computing domain, parallel jobs are of two kinds and those are dependent and independent. We must be cautious during performing the former kind. That job contains some issues related to communication. There independent jobs must be performed by using many VMs at a time.

**Reduction of Cost and Maximizing of Resources:** It is must to manage the constraints which meets the resource allocation related with the cloud operating costs and also to increase the resource usage. In contrast, service provider should provide the services to users with less cost.

**Maintaining High Availability:** The resource availability in cloud is to be assured regarding job with long time evaluations. So some of the techniques must be instantly controls any interrupts or resource unavailability and shifts the jobs to the available resource. Those techniques should also satisfy the property of transparency.

## 4. Classification of Cloud Resources

Cloud computing provides an environment in which the resources are requested for providing services to cloud users by internet. Therefore we can make sure that a Cloud provides computing and makes its availability to every user on requirements. Many researchers were classified resources into two categories such as Physical resources and logical resources otherwise software resources or hardware resources [19, 20].In the cloud computing, cloud provider's controls many resources. Since the cloud computing is a utility based computing, our paper work describes the classification of utility based resources in the cloud computing. The classification includes the following:

**1. Fast Computation Utility:** This kind of resources delivers computational utility at greater speed in the cloud computing platform. By the quick computational utility it provides CaaS (Computation as a Service). Quick computation utility includes the processing capacity, memory size, efficient algorithms and so on.

**2. Storage Utility:** It comprises of numerous hard drives, DB servers and so on. Since system are about to fail at any time so the data redundancy is needed here. Because of cloud's time variant service model, the storage utility must provide some features such as cloud elasticity. By the storage utility the cloud computing delivers SaaS (Storage as a Service).

**3. Communication Utility:** It is otherwise known as Network utility or Naas (Network as a Service). There is no thought about fast computationutility and storage utility without communication utility. The communication utility includes with physical and logical resources. In the cloud computing every service is delivered via high speed internet.

**4. Power / Energy Utility:** In todays the researchers are striving upon research work related to energy efficient techniques in the cloud computing. The energy cost is highly minimized by applying power related approaches. Because of numerous data servers the consumption of power is more in the cloud computing. Cooling gadgets and UPS are considered as the secondary resources. 5. Security Utility: Security will be the biggest issue all the time in any computing domain. Being a user of cloud we required more reliability, trustworthiness, safe and secured services from a cloud.

### 4.1 Cloud Resource Management

In the cloud computing, the main objective of resource management is providing the high availability and sharing of resources, obtaining the time variant service model, providing the utilization of resources in reliable and efficient manner [22]. In perspective with cloud computing, the process which handles all the above stated resources in efficient way along with the assurance to cloud users about QoS is said as Resource management. Under this section the Taxonomy of Cloud resource management is described and is shown in Figure 2. We are presenting the taxonomy as total sequential process in the two phases.

**Phase 1**

In the initial phase the allocation of resources is done in a way that those are requested by application (instead by cloud users) firstly. To complete this phase there are different sequential steps to be considered and it is represented in Figure 3.

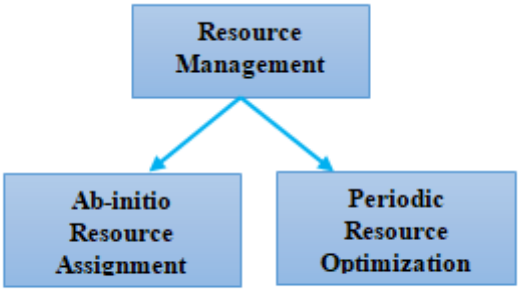


**Figure 2: Taxonomy on Resource Management in Cloud.**

**1. Request Identification**: Initially the requests are identified and it is the initial step of Ab-initio Resource Assignment. Here many resources are identified by the cloud providers. **2. Resource Gathering / Resource Formation**: The very next step to request identification is gathering or forming of the resources is done. In this step the available resources are identified. Custom resources might also be prepared in this step. **3. Resource Brokering**: The resources are negotiated with cloud users under this step in order to confirm that the resource availability is provided on demand basis.**4. Resource Discovery:** In this step, the integration of many resources is done logically based on cloud user's demand. **5. Resource Selection:** Selecting the best resource from the available resources is done under this step. This happens based on the cloud user's requirements.

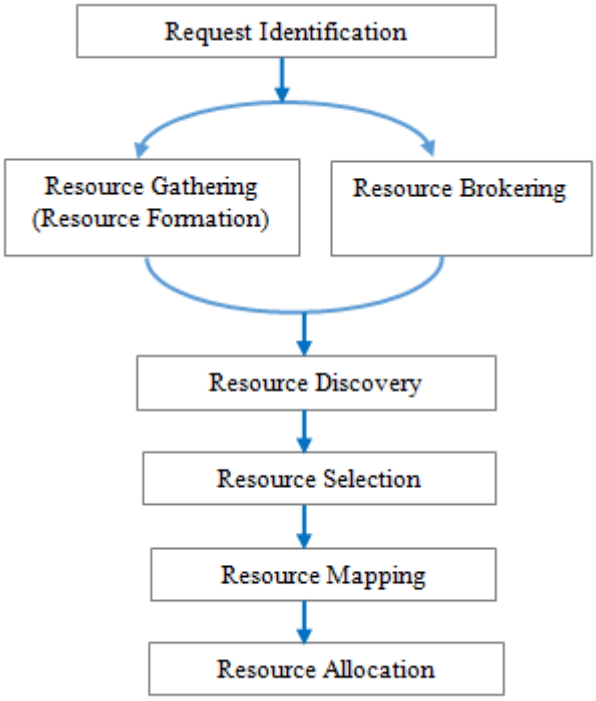


**Figure 3: Ab-initio Resource Assignment**

**6. Resource Mapping: Mapping of** virtual resources with the physical resources (such as a node, a link and so on) is done here.

**7. Resource Allocation:** In this step, the resources are allocated to cloud users. Its key objective is to meet the needs of cloud users and generates profits for the cloud providers.

**Phase 2**

The periodic resource optimization is represented as a procedure for the two kinds of resources such as non-virtualized resources and virtualized resources. The former kind of resources is otherwise known as physical resources. The periodic resource optimization steps are same for both the types of resources. One difference is that the virtualized resources can be assembled and disassembled together based on resource requirement. Therefore, periodic resource optimization for the virtualized resources involves two steps (Resource Bundling and Resource Fragmentation) additionally than to the non-virtualized resources.

**1. For Non-virtualized Resources:** *(a) Resource Monitoring***:** The first and foremost important step in the Periodic Resource Optimization is Resource monitoring. Many non-virtualized cloud resources gets monitored for analysing the resource utilization. The major issue with the monitoring of cloud resources is finding and defining the parameters to it. *(b) Resource Modelling / Resource Prediction*: Under this step, many non-virtualized resources will be predicted by the cloud user's applications. It is said as one of the complex steps since the cloud

resources are not having the same features. *(c) Resource Brokering*: The non-virtualized resources are negotiated with cloud users under this step in order to confirm that the resource availability is provided on demand basis. *(d) Resource Adaptation:* Based on the cloud user's need, the non-virtualized resources in cloud are done scaling. In this step there might be a chance of increase in cost. *(e) Resource Reallocation***:** In this step the reallocation of resources is done to cloud users. The main aim of this is satisfying the requirements of a cloud user and generating profits to the cloud providers. *(f) Resource Pricing:* In the cloud users and providers perception, resource pricing is the major step. Pricing is done based on how the resource of a cloud is used.

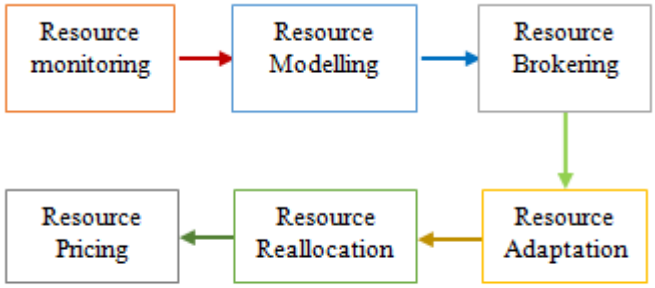


**Figure 4: Periodic Resource Optimization (For non-virtualized Resources).**

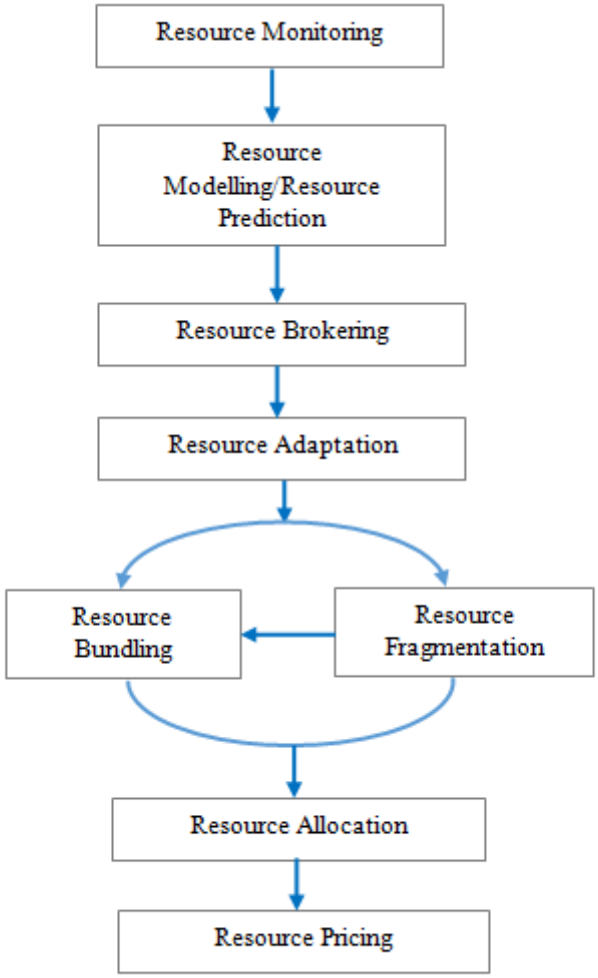


**Figure 5: Periodic Resource Optimization (For virtualized Resources).**

**2. For Virtualized Resources:** *(a) Resource Monitoring*: The first and foremost important step in the Periodic Resource Optimization is Resource monitoring. Many non-virtualized cloud resources get monitored for analysing the resource utilization. The major issue with the monitoring of cloud resources is finding and defining the parameters to it**.** *(b) Resource Modelling / Resource Prediction:* Under this step, many non-virtualized resources will be predicted by the cloud user's applications. It is said as one of the complex steps since the cloud resources are not having the same features**.** *(c) Resource Brokering*: The non-virtualized resources are negotiated with cloud users under this step in order to confirm that the resource availability is provided on demand basis. *(d) Resource Adaptation:*Based on the cloud user's need, the non-virtualized resources in cloud are done scaling. In this step there might be a chance of increase in cost**.** *(e) Resource Bundling*: Based on the requirements many of the non-virtualized resources are bundled into the virtualized resources. *(f) Resource Fragmentation*: Many virtualized resources must be fragmented to bundle them to non-virtualized resources free. Next to this step many non-virtualized resources are bundled in to the virtualized resources as a part of resource bundling. *(g) Resource Reallocation:* In this step the reallocation of resources is done to cloud users. The main aim of this is satisfying the requirements of a cloud user and generating profits to the cloud providers. *(h)*

*Resource Pricing*: In the cloud users and providers perception, resource pricing is the major step. Pricing is done based on how the resource of a cloud is used.

## 5. Conclusion

Resource management is the foremost concern during providing the demanded resources and also computation in the cloud environment. This paper presents many research concerns related to cloud resources management by comparing with the prevailing resource allocation systems. We discussed some of the problems and challenges such as resource assignment, scheduling of jobs, balancing of loads, scalability, pricing, availability and energy management in this paper. We also described the components of comprised in cloud resource management system. Besides we surveyed the research on load balancing, allocation of resources and workflow scheduling. It leads to attain complete knowledge on resource management algorithms. We also presented the description about the taxonomy of the cloud resources in dual phases such as for non-virtualized resources and for virtualized resources. This survey helps in filling the gap that exists in between the work done and the remaining to be done.

## References

[1]. https://searchcloudcomputing.techtarget.com/definition/cloudcomputing

[2]. Sajjan R.S, Biradar Rekha Yashwantra, "Load Balancing and its Algorithms in Cloud Computing: A Survey", International Journal of Computer Sciences and Engineering, Volume-5, Issue-1, Jan. 2017.

[3]. Fuhui Wu, Qingbo Wu, Yusong Tan, "Workflow scheduling in cloud: a survey", Journal of Supercomputing, May 2015.

[4]. S. S. Manvi and G. K. Shyam. Resource management for infrastructure as a service (iaas) in cloud computing: A survey. Journal of Network and Computer Applications, 41:424–440, 2014.

[5]. S. M. Parikh. A survey on cloud computing resource allocation techniques. In Engineering (NUiCONE), 2013 Nirma University International Conference on, pages 1–5. IEEE, 2013.

[6]. L. Xu and J. Li. Building efficient resource management systems in the cloud: Opportunities and challenges. International Journal of Grid and Distributed Computing, 9(3):157–172, 2016.

[7]. Sivadon Chaisiri, Bu-Sung Lee, Dusit Niyato, "Optimization of Resource Provisioning Cost in Cloud Computing", IEEE Transactions on Services Computing, Volume: 5, April-June 2012.

[8]. Seyedehmehrnaz Mireslami, Logan Rakai, Behrouz Homayoun Far, Mea Wang, "Simultaneous Cost and QoS Optimization for Cloud Resource Allocation", IEEE Transactions on Network and Service Management, Volume: 14, Issue:3, Sept. 2017.

[9]. Dilip Kumar S. M., Naidila Sadashiv, R. S. Goudar, "Priority based resource allocation and demand based pricing model in peer-to-peer clouds", International Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 2014.

[10]. Aarti Singh, Dimple Juneja, Manisha Malhotra, "A novel agent based autonomous and service composition framework for cost optimization of resource provisioning in cloud computing", Journal of King Saud University - Computer and Information Sciences, Volume 29, Issue 1, January 2017.

[11]. Moona Yakhchi, Seyed Mohssen Ghafari, Shahpar Yakhchi, Mahdi Fazeli, Ahmad Patooghi, "Proposing a Load Balancing Method Based on Cuckoo Optimization Algorithm for Energy Management in Cloud Computing Infrastructures", 6th International Conference on Modeling, Simulation, and Applied Optimization (ICMSAO), May 2015.

[12]. S. Keshvadi, B. Faghih, "A multi-agent based load balancing system in IaaS cloud environment", International Robotics & Automation Journal, Oct. 2016.

[13]. D. Komarasamy, V. Muthuswamy, "A novel approach for dynamic load balancing with eff ective Bin packing and VM reconfiguration in cloud", Indian Journal of Science and Technology, Vol 9, March 2016.

[14]. X. Wei, J. Fan, T. Wang, Q. Wang, "Efficient application scheduling in mobile cloud computing based on MAX–MIN ant system", Soft Computing, Volume 20, Issue 7, July 2016.

[15]. Weiwei Lin, Chen Liang, James Z. Wang, Rajkumar Buyya, "Bandwidth - aware divisible task scheduling for cloud computing", Journal of Software Practice and Experience, Vol. 44, Feb. 2014

[16]. Saurabh Kumar Garg, Adel Nadjaran Toosi, Srinivasa K.Gopalaiyengar, Rajkumar Buyya, "SLA-based virtual machine management for heterogeneous workloads in a cloud datacenter", Journal of Network and Computer Applications, Volume 45, October 2014.

[17]. Sahar Adabi, Ali Movaghar, Amir Masoud Rahmani, "Bi-level fuzzy based advanced reservation of Cloud workflow applications on distributed Grid resources", The Journal of Supercomputing, Volume 67, Issue 1, January 2014.

[18]. Upendra Bhoi, Purvi N. Ramanuj, "Enhanced Max-min Task Scheduling Algorithm in Cloud Computing", International Journal of Application or Innovation in Engineering & Management, Volume 2, Issue 4, April 2013

[19]. B. Jennings and R. Stadler. Resource management in clouds: Survey and research challenges. Journal of Network and Systems Management, 23(3):567–619, 2015.

[20]. S. Singh and I. Chana. Qos-aware autonomic resource management in cloud computing: A systematic review. ACM Computing Surveys (CSUR), 48(3):42, 2015.

[21]. S. S. Manvi and G. K. Shyam. Resource management for infrastructure as a service (IaaS) in cloud computing: A survey. Journal of Network and Computer Applications, 41:424–440, 2014.

[22]. L. Xu and J. Li. Building efficient resource management systems in the cloud: Opportunities and challenges. International Journal of Grid and Distributed Computing, 9(3):15

**Author Biography:**

**Dr. G. Rama Subba Reddy** as an Associate Professor and Head of the Department at Computer Science & Engineering, Mother Theresa Institute of Engineering & Technology. Palamaner, Andhra Pradesh. He is the author of over 50 articles in various fields like cloud computing, IoT, WSN, Network protocols & Security. His current research focus is on Machine Learning, Data mining, and Cloud computing.