American Journal of Computer Technology and Application
Vol. 2, No. 4, April 2014, PP: 1 - 9, ISSN: 2327-2325 (Online)
Available online at www.ajcta.com

Research article

Optimization Distributing Resources to Create an Efficient Cloud Computing

Abbass Ali Alizadehbarmi [1] , Seyed Yaser Bozorgi Rad[2] , Mehdi Rahimi Amirbandeh[3]

[1] faculty of Computer Science, Islamic Azad University- Babol branch, IRAN
[2] faculty of Computer Science, Islamic Azad University- Babol branch, IRAN
[3] *Ayandegan Institute of Higher Education, Tonekabon, Iran*
* alizadehbarmi@yahoo.com
* y.bozorgi.r@gmail.com
*rahimi_1360_mehdi@yahoo.com

Abstract:

Increasing demand for most efficient computing and data storage in distributed computing systems in recent years has attracted the attention of many researchers.

The aim of this paper is to investigate how resource allocation in cloud computing can be used for creating better performance in this area. There are many challenges and problems in resource allocation. The most important of these challenges is the management of resources. Given the challenges and problems in the process of resource allocation in cloud computing ،by categorizing existing problems in this issue in a step-by-step way for eliminating them, and this solution is an ideal approach to optimize allocations of dynamic resources and also to improve clouding computing.

In this regard, with detailed analysis of Run-time performance management and utilizing distributed data in order to increase the ability of managing distributed data for greater access to the resources to optimize resource allocation in cloud computing, we propose a model.

The results of the proposed model for optimizing the performance of resource allocation are presented in the form of stimulation and the results of stimulation indicate the optimized and desirable nature of the proposed model.

Keywords: Cloud Computing, Distributing Resources, request runtime

---

1. Introduction:

Cloud computing [1] is a common terminology used in both business and academic fields [7]. In recent years, Cloud Computing has become an emerging technology that gains wide influence on IT systems. Cloud Computing is a distributed computing model for enabling service-oriented, on-demand network access to rapidly scalable resources [9]. Such resources include infrastructure as a service (IaaS), development and runtime platforms as a service (PaaS), and software and business applications as a service (SaaS).[5] Iaas is very similar to a private cloud, except for the fact that clients do not own the server. Instead, a third party allows client to install client‘s own virtual server on their IT infrastructure in exchange for a rental fee [4].
Clients do not own the resources, yet applications and data are guaranteed to be available and ubiquitously accessible by means of Web services and Web APIs “in the Cloud”.[5] Clients in this system are application software that require processing, data storage and communication resources in this “on-demand capacity provisioning” or “lease model of the IT infrastructure” [3].
Cloud computing is a combination of several concepts including virtualization, resource pooling, resource monitoring, dynamic provisioning, utility computing, multi-tenancy, and elasticity. The main entities of cloud computing are service providers, physical resources, virtualized resources, and end-users [7].
Resource management in distributed systems is one of the most important challenges. Due to more requests for resource in these systems, input size of this problems are much larger than ordinary resource allocation problems in centralized computing systems. There is number of papers discussing the resource allocation in grid computing systems e.g.,.[3] Due to difference between cloud computing system and grid computing system, it is necessary to address the resource allocation problem in cloud computing systems separately.[3]
The main objective is to distribute the resources that users and applications to easily access remote resources. And it controlled and effective ways to share. Resources can be anything. Such as printers, computers, storage, data, files, web pages and networks [11].
In this paper, we examine the management of run-time performance and the use of distributed data for efficient management to gain greater access to resources, and then to optimize resource allocation in cloud computing. To achieve the goal (i) Review and evaluation of the request runtime functions (ii) Distributed data types and how to use distributed data to increase the ability of management. With the help of the above objectives, we will propose a conceptual model to achieve the desired goal.
In the following sections, further analysis of the proposed model, simulation results, conclusion, and further works to be done will be presented.

2. Review and evaluation of the request runtime functions

The distributed cloud provides scalability and reduces the limitations of the network, while all resources are fully decentralized. Cloud distributed can contain billions of computing nodes resources to the user's machine and is much less expensive. Each node knows that it is only a subset of all the nodes in the cloud. To perform certain calculations, a user requires a certain amount of resources (CPU power and memory). Many users who tend to perform different calculations, the source node can perform calculations to explore and ask questions by many users optimal allocation of node resources to serve the maximum number of users has.
Different cloud providers are using different programming framework for the cloud itself is often used. Now all cloud providers rely on large data centers. Mainly centralized and distributed computing. In distributed computing environment, there are many distributed machines around the world. A central server that contains information about all the machines hosting plan there. In the cloud computing environment, we need a description of the specific resources that can be dedicated to identify resources that allows users to request resources are clearly being used. This describes a distributed computer resources for resource discovery and resource allocation is required.
We examine the implementation of a distributed cloud application resource discovery, Cloud architecture and size of the client application process.

2.1 The discovery of a distributed computer resources

Of using different algorithms to solve this problem, Like the Kademlia algorithm uses the routing nodes or the discovery has XoR done. Mixed Tables or hash: one to find the location of nodes and one for resource discovery in distributed cloud [7].

The service providers should have a service level agreement SLA. These include computing power, storage space, bandwidth, availability and security [3]. If these sources too may be provided to the customer Increased costs incurred in terms of electrical energy costs and carbon emissions of data centers increases.

2.2 Another method of implementation of customer requests

customer classification, clients in each cluster are deciding on a course that helps to compensate for dynamic changes in systems that do not require a decision on the cloud. It can be used in cluster or server. To optimize the performance of the application, Online Measurement of the methods of traditional resource management systems not adopt ways of allocating resources based systems that focus on optimizing the overall performance of the cluster is the cluster approach is now to maximize the system's overall performance and clustering is done. Their goal is to increase throughput and reduce average waiting time of a cluster of CPU utilization and response time for a request to be efficient.
All requests for user do not have equal importance and therefore the actual levels of the service to different users are required.
Other factors contributing to the implementation of the request: Energy efficiency, security and Schedule run the program. Scheduler is responsible for assigning each task executes a resource for the general user is based on quality of service parameters and costs for service providers.
Each time the user request, the cloud providers are shorter, Resources and resource management System that resulted in staff productivity and minimizing SLA violations, and service providers to improve profitability.

3. Distributed data types and how to use distributed data

Cloud computing systems are parts of distributed systems, Collection of independent computers that the users' perspective as a single and coherent system operates. An important feature is that the differences between computers bonding between them, often hidden from the user's perspective. Distributed systems usually are layered software organization. Distributed systems, Often Complex parts that are, by definition, their components are distributed across multiple machines .To restrain the complexity of these systems must be properly organized. There is methods different for organizing distributed systems, But the most obvious, logical organization deciding between a set of software components on the one hand and on the other hand, is a physical system.
Distribution of resources in cloud computing architectures that are used for this system was to be done, the general architecture of cloud computing, there are two parts. One area that the Front end Back end user or customer and the other is that the cloud system. The software interface between a web-based email services such as applications, Web browsers like Internet Explorer, Firefox and Other systems in the network application available to customers there.

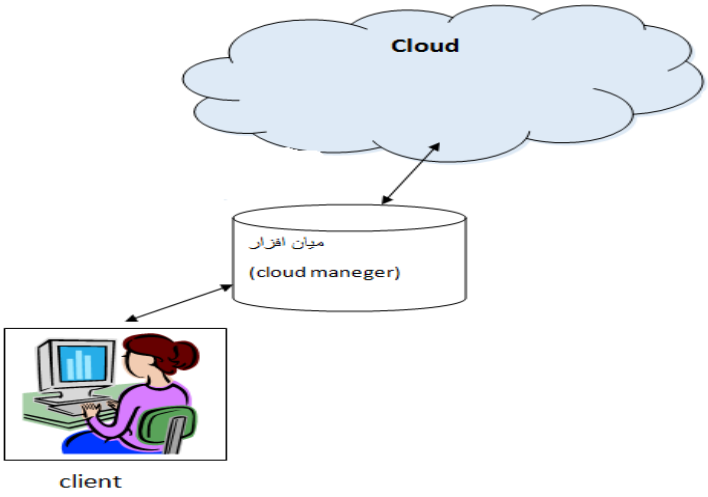

Figure 1: A simplified view of cloud systems architecture

Back end of the cloud computing system, Computers different, Servers and data storage systems that create the cloud computing services. The theory a cloud computing system, the application of data processing is thought to play in the cloud environment. Each has its own dedicated server.
How automate server deployment in the cloud is this means that Different set of machines that can be used at any time. However, each part of the program should be able to call other components in a known address. [5]
One of possible architecture for cloud computing systems proposed in [8] which in that the architecture considered in resource management and scheduling based on a defined response to a user request is made. This architecture is designed based on Manjrasoft Aneka software, [8] which is a software program to build and deploy distributed applications on the cloud. [8]

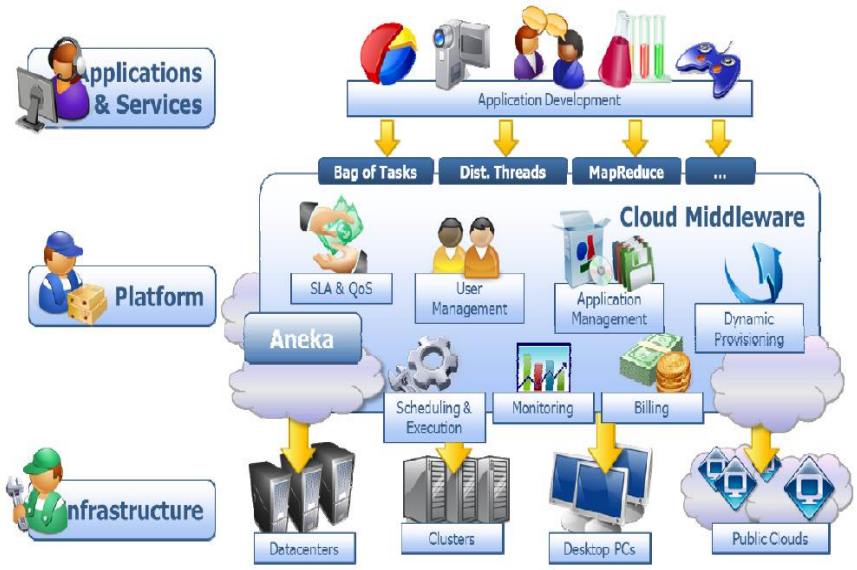

Figure 2: Architecture Aneka [8]

Drtrahy Aneka software system architecture using several key institutions involved.
• Resource Allocation SLA
• Autonomous resource management
• Virtual machines or VMS
In this architecture, services are completely isolated from each other. In fact, the relationship between services is through sending messages. Because the services are separated if for some reason one of the leading service Aneka service is not responding, other services to be regularly maintained.

Services of Aneka implement with expanding basic services. There are two important points here; one is expansion of basic services and the other is to defining messages to enable applications SLA. In this architecture, dynamically provisioning process for autonomous SLA. This architecture uses a scheduler, service you will receive a new job. Due to resource estimates that do or ask how long it takes however, this estimate may be inaccurate. Scheduler checks if some of the resources allocated resources can be harvested and placed in other jobs available.
If resources are not sufficient for a job under provisioned case goes to work and if time is achievable with the resources available to the state is provided. Finally, if one cannot estimate the time the work will go into the raw mode. In each of the cases, a message is sent to the Scheduler service to service. Status is under provisioned Message containing the number of resources needed to be worked.

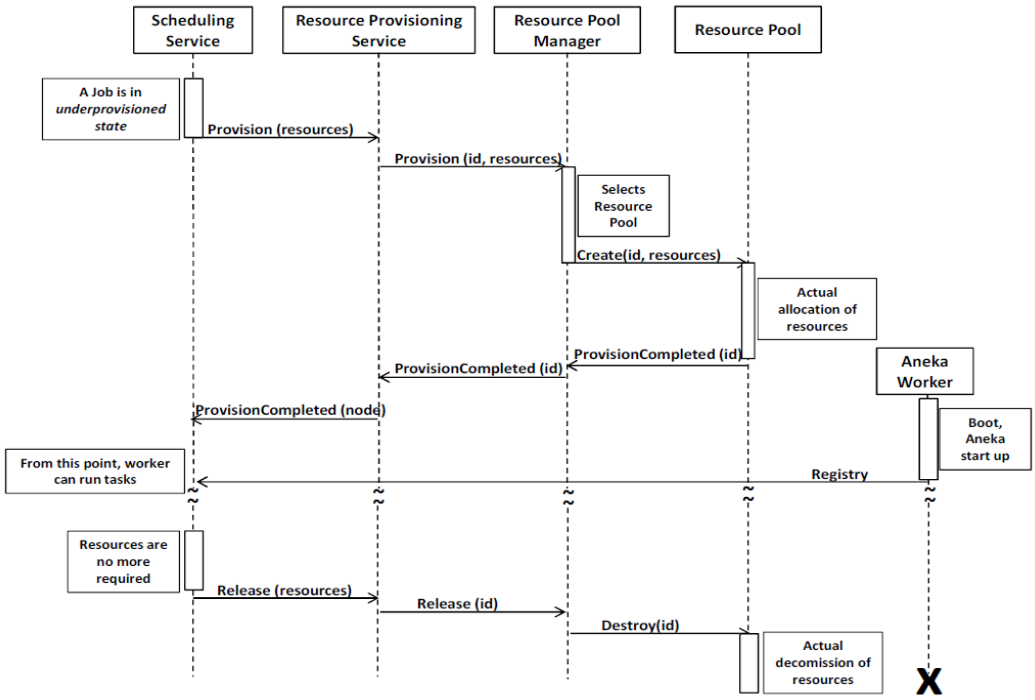

Figure 3: Interaction between Aneka services provided during the process dynamics [8]

4. The proposed model

The first step is to optimize resource management in the proposed model, the user request in the least possible time.

4,1.Step 1: routing request a cloud computing system

As the in the beginning paper was raised. Cloud computing system is composed of billions of nodes so the first step towards the implementation of a cloud computing system is the routing of user requests Ask the users / customers the nearest server that is best for your request to be sent. For this work, we propose to use intelligent search algorithms that in this algorithm, the objective function is to find the closest server users to run the application in the shortest possible time. And provide resources to implement user requests. Our proposed algorithm is a search algorithm for the job. The algorithm uses multiple paths instead of the current state of the route uses. This algorithm is chosen because of its advantages: 1 - uses very little memory. 2 - For the great and infinite (such as cloud computing system) is the perfect solution. This is because the search for algorithm based on target function at the beginning of this section is raised, be the algorithm for optimization problems such as resource management is a very useful and helpful.
Piece of code to implement this algorithm in finding the optimal solution for users to find the best server to run the application:

4,1,1. Runtime architecture data distribution in the model

After selecting the best server in the continued work to implement user requests, Turns to the user application is running. The model we Aneka from the fusion of architecture, which is described in the previous section sliding window method is to implement user requests.
After the user request to server (services) sent. It is possible that at the time of server resources available to other users that It all server resources it is not required. So this leads to wasted resources and the SLA is violated.

4,2. Step 2: Determine scheduling and turn off the source of non-urgent

Our first suggestion is based on the architecture of Aneka That one Schedule a service provider (server) exist. One Deadline Be created for each request and application server in a queue with priority placed Requests based on priority (Which admittedly are not the same priority) Server resources needed at a given time for any request for the user puts. For the server, all the resources available for implementing user requests not to work, The only resources that are required to be provided And resources necessary to disable or silence mode and other users should be provided when required. This result is to manage resources and avoid wasting resources.

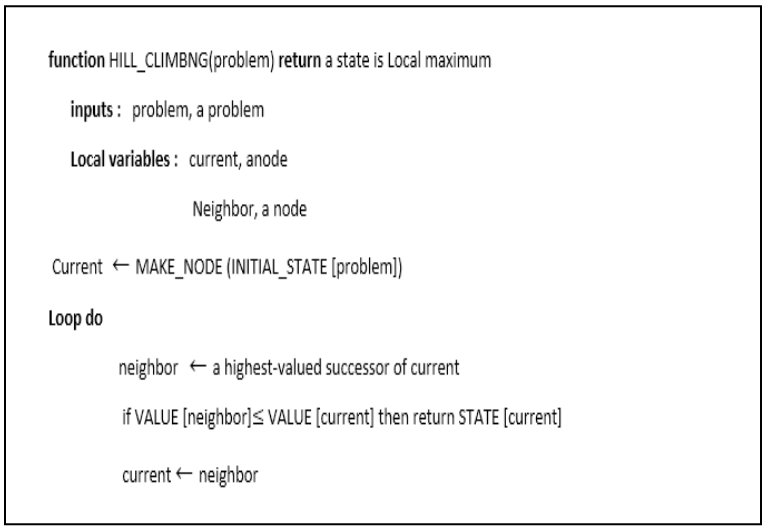

Figure 4: A view the proposed algorithm targets based on the proposed solution[10]

4,3. Step 3: Declaration Resources needed

Here is a question that is that: How Server resources requirements User to run request to be notified?
The answer to this question can be raised species that, when a user sends his request. Request to the resource manager (the firmware) arrives.

4,4.. Step 4: How to run a query between the middleware and the service provider

Now the resource manager sending messages frames in the sliding window method. Resources needed to run the application server proclaims. The server received the message. And in a given time period apply for it only resources that are needed has enabled and other sources will turn off. The user requested to be packaged in firmware brings.

4,5. Step 5: Send a request in frame

These packets are placed in a series of Seq #. Due to the size of the frame (window) Requirements Packet n put it inside. Firmware requests in the form the frame is sent to the processing system. And waits an Ack message the user receives packets the has received then the next frame to be sent.

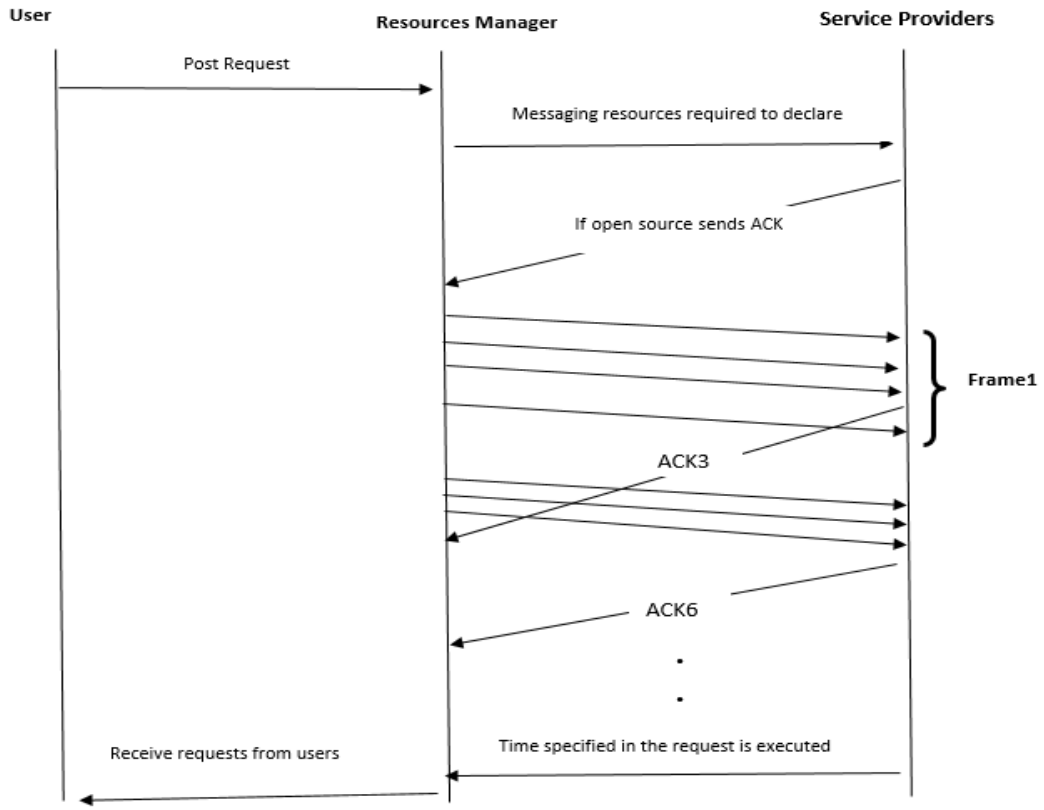

Figure 5: Run the application

If the message primary submitted and resources involved and server resources can give priority should be placed in a queue. However, due to the time that is given the job done. At this Conditions There may be different modes:
1 - Run the application schedule specified by the service provider to an end.
2 - Run Application do not schedule (Job failed)
3 - The job is canceled by user request.

5. Simulation results

To evaluate the proposed solution, Resource allocation, MATLAB software framework and 2 Cloudsim3 are used for simulation.
Here we offer two charts:
Figure 1 routing solution using informed heuristics to find best server for executing user requests in the shortest time possible is. In this section we Machines between the average capacity of 3.5 GB RAM and 300 GB storage capacity and is used for 64-bit platforms are Linux-based machines.

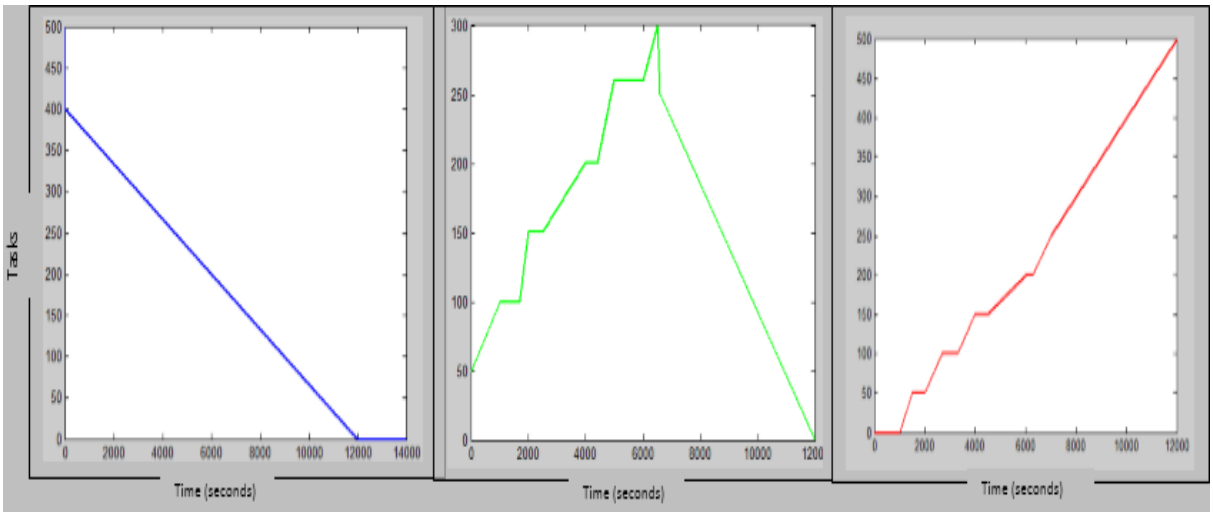

Figure 2: Diagram of things done in the required

And Figure 2, the graph of doing things Time is determined by the three modes are displayed on the chart.
In both graphs runtime (Time) has been considered. Simulation results show the efficiency of the proposed model is useful for creating cloud computing.
In this simulation, it has been assumed that the first server closest server to provide services the resources will be available in the shortest possible time. Here, the distribution of resources in resource distribution model is proposed.

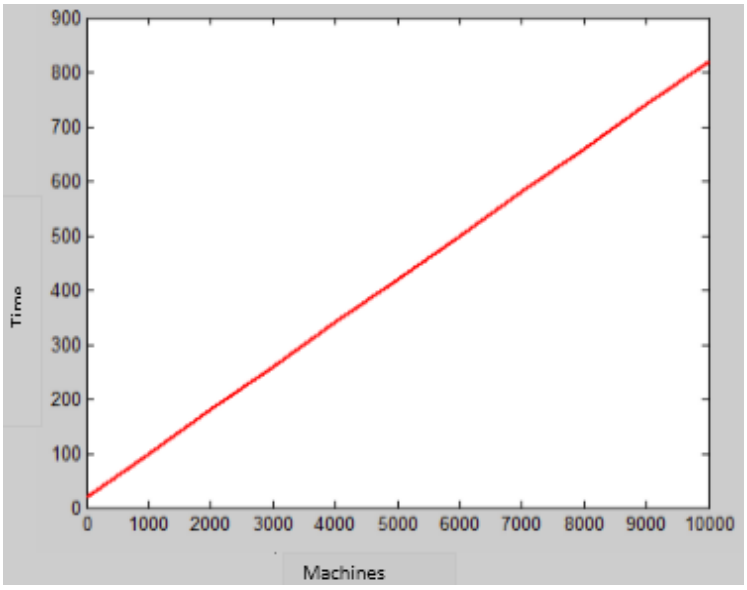

Figure 1: The routing solution using heuristic conscious

6. Conclusions

In this paper, our goal is to explore challenges existing in the SLA to create a cloud-based resource allocation is efficient. Examining the current challenges in the field of resource management is the most critical challenges in considering the application execution time Members and the type of data distribution (data architecture) in a cloud computing system; we paid to get the problem resolved.
However, there is still room for further work in this area. In the future it can be used on most of the time of the application software decrease in the effective cloud computing system is optimized.
We think that more comprehensive investigations are needed in this area to enable better resource management for distributed data fields. It results in greater profitability for service providers.

REFERENCES:

[1] David Villegas, Athanasios Antoniou, Seyed Masoud Sadjadi, and Alexandru Iosup (2011)," An Analysis of Provisioning and Allocation Policies for Infrastructure-as-a-Service Clouds: Extended Results", Draft

[2] Guillaume Pierre, Thilo Kielmann, Emanuele Rocca…(2013), " ConPaaS: an Integrated Runtime Environment for Elastic Cloud Applications", HPDC 2013

[3] Hadi Goudarzi and Massoud Pedram(2010)," Maximizing Profit in Cloud Computing System via Resource Allocation", University of Southern California, Los Angele

[4] Jasnil Bodele, Anil K. Sarje (2013)," DyanamicLoad Balancing With Cost and Energy Optimization In Cloud Computing", International Journal of Engineering Research & Technology (IJERT)

[5] Leitner, P.Distrib. Syst. Group, Vienna Univ. of Technol., Vienna, Austria Hummer, W.; Satzger, B.; Inzinger, C.; Dustdar, S.(2012), "Cost-Efficient and Application SLA-Aware Client Side Request Scheduling in an Infrastructure-as-a-Service Cloud",

[6] Michael Hauck, Matthias Huber, Markus Klems, Samuel Kounev,J¨ orn M¨uller-Quade, Alexander Pretschner, Ralf Reussner, Stefan Tai," Challenges and Opportunities of Cloud Computing",Karlsruhe Institute of Technology

[7] Praveen Khethavath, Johnson Thomas, Eric Chan-Tin and Hong Liu(2011)," Introducing a Distributed Cloud Architecture with Efficient Resource Discovery and Optimal Resource Allocation", Computer Science,

[8] Rajkumar Buyya, Saurabh Kumar Garg, and Rodrigo N. Calheiros(2011)," SLA-Oriented Resource Provisioning for Cloud n Computing: Challenges, Architecture, and Solutions", International Conference on Cloud and Service Computing

[9] Rajkumar Buyya, Rodrigo N. Calheiros, and Xiaorong Li (2012)," Autonomic Cloud Computing: Open Challenges and Architectural Elements", Cloud Computing and Distributed Systems (CLOUDS) Laboratory

[10] Russell, Stuart Jonathan (1395), "Approach New at Artificial intelligence", InTech. pp. 129-138

[11] tanenbaum,andrew s,steen,maarten van (2007),'distributed system:princeples and paradigms 2nd ed.c,in tach pp.10-50