BINARY PSOGSA FOR LOAD BALANCING USING TASK SCHEDULING IN CLOUD ENVIRONMENT

Keyur Patel[1], Nimisha Patel[2],
Sankalchand Patel College of Engineering, Sanakalchand Patel University
Visnagar-384315, State: Gujarat Country: India

Abstract –

In cloud environments, load balancing task scheduling is an important issue that directly affects resource utilization. Unquestionably, load balancing scheduling is a serious aspect that must be considered in the cloud research field due to the significant impact on both the back end and front end. Whenever an effective load balance has been achieved in the cloud then good resource utilization will also be achieved. An effective load balance means distributing the submitted workload over cloud VMs in a balanced way, leading to high resource utilization and high user satisfaction. In this paper, we propose a load balancing algorithm, Binary Load Balancing – Hybrid Particle Swarm Optimization and Gravitational Search Algorithm (Bin-LB-PSOGSA), which is a bio-inspired load balancing scheduling algorithm that efficiently enables the scheduling process to improve load balance level on VMs. The proposed algorithm finds the best Task-to-Virtual machine mapping that is influenced by the length of submitted workload and VM processing speed.

Keywords—Cloud Computing; load balancing; task scheduling; Data Security

1. INTRODUCTION

Services that supply an on-demand environment for cloud users such as developing, testing, delivering, and managing software applications. It is mainly used by application and software developers. Third and finally, the service model SaaS delivers software applications over the Internet to cloud users on-demand and typically on a subscription basis. It is essential to cloud providers to tend the management operations in both task-level and resource-level services. The task-level scheduling allocates a task to a virtual machine (which we address in this study), while the resource-level scheduling allocates a virtual machine to a host. The other important issue is to keep cloud resources balanced. Therefore, they also tend to schedule the incoming application requests to virtual machines in order to complete submitted tasks at the expected time in a balanced way. Numerous objectives have been addressed in the literature, such as minimizing make span, maximizing load balancing, minimizing flowtime, and minimizing monetary cost.

.

Cloud services allow individuals and businesses to use software and hardware that are managed by third parties at remote locations. Examples of cloud services include online file storage, social networking sites, webmail, and online business applications. The cloud computing model allows access to information and computer resources from anywhere that a network connection is available. Cloud computing provides a shared pool of resources, including data storage space, networks, computer processing power, and specialized corporate and user applications [1]. Cloud services are popular because they can reduce the cost and complexity of owning and operating computers and networks. Since cloud users do not have to invest in information technology infrastructure, purchase hardware, or buy software license, the benefits are low upfront costs, rapid return on investment, rapid deployment, customization, flexible use, and solutions that can make use of new innovations. In addition, cloud providers that have specialized in a particular area (such as e-mail) can bring advanced services that a single company might not be able to afford or develop.

Some other benefits to users include scalability, reliability, and efficiency. Scalability means that cloud computing offers unlimited processing and storage capacity. The cloud is reliable in that it enables access to applications and documents anywhere in the world via the Internet. Cloud computing is often considered efficient because it allows organizations to free up resources to focus on innovation and product development. Another potential benefit is that personal information may be better protected in the cloud. Specifically, cloud computing may improve efforts to build privacy protection into technology from the start and the use of better security mechanisms. Cloud computing will enable more flexible IT acquisition and improvements, which may permit adjustments to procedures based on the sensitivity of the data. Widespread use of the cloud may also encourage open standards for cloud computing that will establish baseline data security features common across different services and providers. Cloud computing may also allow for better audit trails .The characteristics of cloud computing tells storage is not done on a single system, User terminal is only responsible for user interaction and access to service, Service can be provided directly to user terminal or access through network, Provides reliable secure data, up-to-date, no virus attacks, Users need not worry on configuration and it is easily manageable, Multitenancy, pay as you use, scalable. `They can improvise cloud computing by virtualization, reduce device dependency, Platform independent, Integrate resources.

2. BACKGROUND THEORY

2.1 Architectural Components

Cloud Services models are commonly divided into SaaS, PaaS and IaaS that exhibited by a given cloud infrastructure. It's helpful to add more structure to the service model stacks: Figure 1. Shows a cloud reference architecture that makes the most important security relevant cloud components explicit and provides an abstract overview of cloud computing for security issue analysis[2].

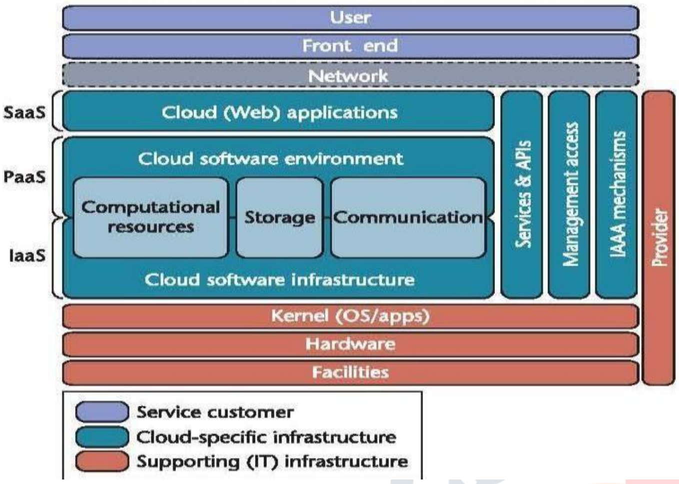


Figure 1: The Cloud reference architecture

Cloud Computing can significantly reduce the cost and complexity of owning and operating computers and networks. If an organization uses a cloud provider, It need not spend money on information technology infrastructure or buy hardware or software licenses.

2.2 Types of Cloud Service

Cloud Computing consists of three different types of service provision as shown in figure 2. In Each case services are hosted remotely and accessed over internet through customer web browser, rather than being installed locally on customer's computer. Firstly SaaS (Software as a Service) refers to the provision of software applications in the cloud. Secondly PaaS (Platform as a Service) refers to the provision of services that enable the customers to deploy in the cloud, application created using programming languages and tools provided by the supplier. Thirdly IaaS (Infrastructure as a Service) refers to the services providing computer Processing power, storage space and network capacity, which enable the customers to run arbitrary software in the cloud.

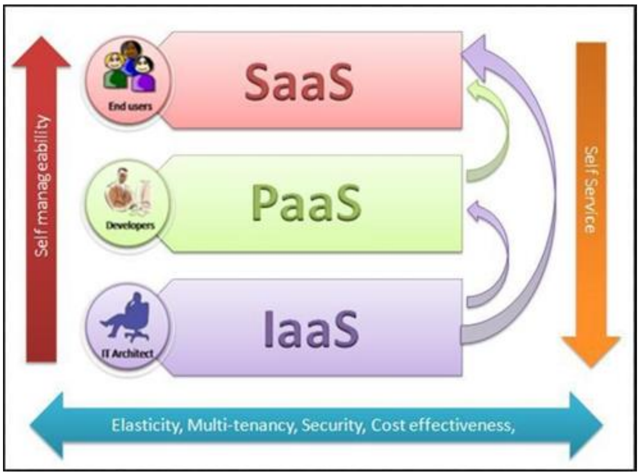


Figure 2: SaaS, PaaS, IaaS [1]

Software as a Service (SaaS):

Cloud consumers release their application in a hosting environment, which can be accessed through networks from various clients (e.g. Web browser, PDA, etc.) by application user. Cloud consumers do not have control over the cloud infrastructure that often employs multi-tenancy system architecture, namely, different cloud consumers applications are organized in a single logical environment in the SaaS cloud to achieve economies of scale optimization in terms of speed, security, availability, disaster recovery and maintenance. Examples of SaaS include SalesForce.com, Google Mail, Google Docs and so forth [2].

Platform as a Service (PaaS):

PaaS is a development platform supporting the full "Software Lifecycle" which allows cloud consumers to develop cloud services and applications (e.g. SaaS) directly on the PaaS Cloud. Hence, the difference between SaaS and PaaS is that SaaS only hosts completed cloud applications whereas PaaS offers a development platform that hosts both completed and in-progress cloud application. An Example of PaaS is Google AppEngine [2].

Infrastructure as a Service (IaaS):

Cloud consumers directly use IT infrastructures (processing, storage, networks and other fundamental computing resources) provided in the IaaS cloud. Virtualization is extensively used in IaaS cloud in order to integrate/decompose physical resources in an ad-hoc manner to meet growing or shrinking resource demand from cloud consumers. The basic strategy of virtualization is to set up independent virtual machines that are isolated from both the underlying hardware and other VMs. An Example of IaaS is Amazon's EC2 [2].

2.3 Cloud Formations

The cloud environment is subdivided into public, private, hybrid and community clouds as shown in Figure 3.

Public Cloud: Public Clouds are those in which services are available to the public at large over the internet. Public clouds provide an elastic, cost-effective means to deploy solutions and take care of deploying, managing, and securing the infrastructure. Companies can use it on demand, and with pay-as-you-option, it is much like utility consumption [1].

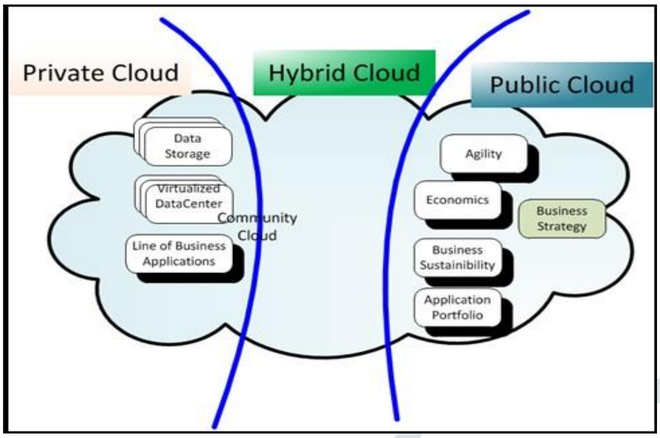


Figure 3: Type of Cloud [1]

Private Cloud: Private Cloud is essentially a private network used by one customer for whom data security and privacy is usually a primary concern. The downside of this type of cloud is that the customer will have to bear the significant cost of setting up and then maintaining the network alone [3].

Hybrid Cloud: Hybrid cloud environments are often used where a customer has requirements for a mixed set of dedicated server and cloud hosting, for example if some of the data being stored is of a very sensitive nature. Another reason for using hybrid clouds is when the organization needs more processing power than is available in-house and obtains extra requirement in the cloud.

2.5 Load Balancing in Cloud Computing:

Cloud computing is a federation of many resources interacting together by sharing and pooling resources efficiently. Day by day millions of data are created every second this leads in increase of web traffic in web world. Load balancing has become imperative and important need of balancing the load on this heavy traffic. Load balancing mechanism can be constructed by many strategies and techniques. There are dedicated servers allocated to handle this load balancing across the servers, if there is failure of any node in network it will redirect the load to the running instance or node. Elastic Load Balancing also performs health checks on each instance. If an instance is not responding, the load balancer will automatically redirect traffic to the healthy instances [4].

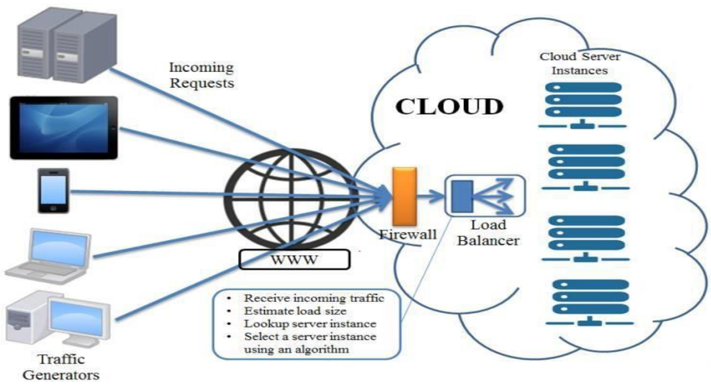


Figure 4: Load Balancing Architecture [4]

Figure 4 depicts load balancing architecture used in cloud environment where load balancer load traffic using following steps [4]: Various incoming client request are received. It checks the load size of incoming request and builds a request queue. The current load of the server is checked in the server pool. It uses load balancing algorithms to select appropriate sever.

2.6 Load Balancing Algorithms Execution:

Load balancing is the technique of distributing the load between various resources in any system. Thus, load requires to be distributed over the resources in cloud-based architecture, so that each resource does almost the equal amount of work at any point of time. Basic requirement is to provide some techniques to balance requests to provide the solution of fast response for request.

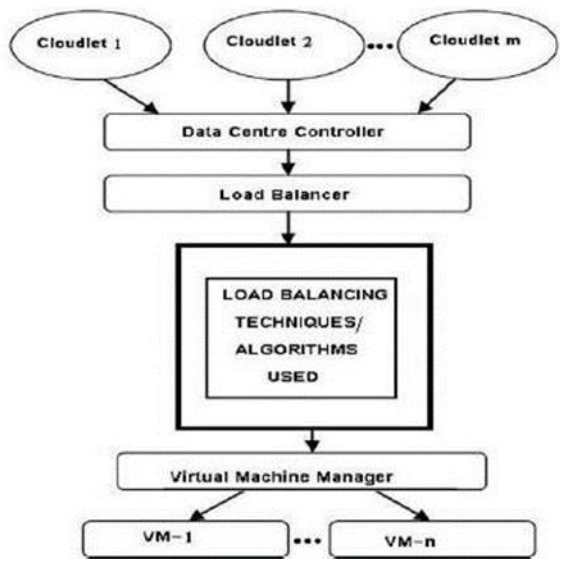


Figure 5: Load Balancing Algorithms Execution [5]

Cloud Load Balancers manage online traffic by distributing workloads between multiple servers and resources automatically. The maximize throughput. Minimize response time, and avoid overload [5].

2.7 Round Robin Algorithm:

In this algorithm the data packet is divided evenly to all processor. The order of process allocation is locally independent. This algorithm works well when the data is equally and number of processor smaller than no. of processor. One of its benefit is that it does not require IPC means inter process communication [4] [6].

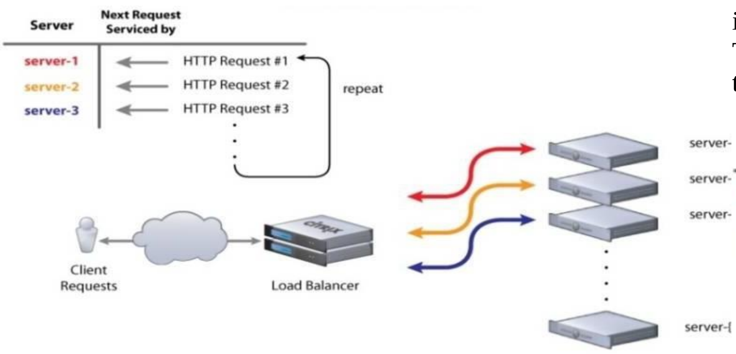


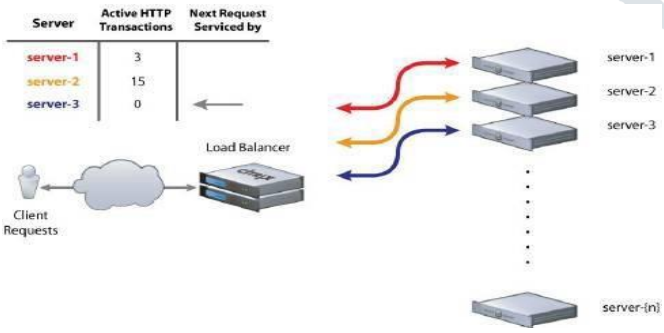
Figure 6: Round Robin Algorithm [4]

This algorithm is used in spatial purpose Application where HTTP request are similar nature and there by distributed equally [6].

The moment a server IP address has been given its IP address is moved to the back of the list of available IP addresses and gradually it moves back to the top of the list and becomes available again. For example, a company has one domain name and three identical copies of the same web site residing on three servers with three different IP addresses. When one user accesses the home page it will be sent to the first IP addresses. The fourth user, therefore, will be sent to the first IP address and so forth [4].

2.8 Weighted Round Robin Algorithm:

This builds on the simple Round Robin load balancing method. In the weighted version each server in the pool is given static numerical weights. Servers with higher ratings get more request sent to them. For example, when a client's start coming in the first 5 will be assigned to server 1 and the 6^{th} to Server 2. If more clients come in, the same sequences will be followed. That is the 7^{th}, 8^{th}, 9^{th}, 10^{th} and 11^{th} will all go to Sewrver1 and the 12^{th} to Server 2



and so on [4].

Figure 7: Weighted Round Robin [4]

2.9 Source IP Hash Algorithm:

Source IP Hash load balancing uses an algorithm that takes the source and destination IP address of the client and server and combines them to generate a unique hash key. This key is used to allocate the client to a particular server. As the key can be regenerated if the session is broken this method of load balancing can ensure that the client request is directed to the same server that it was using previously. This is useful if it's important that a client should connect to a session that is still active after a disconnection [4].

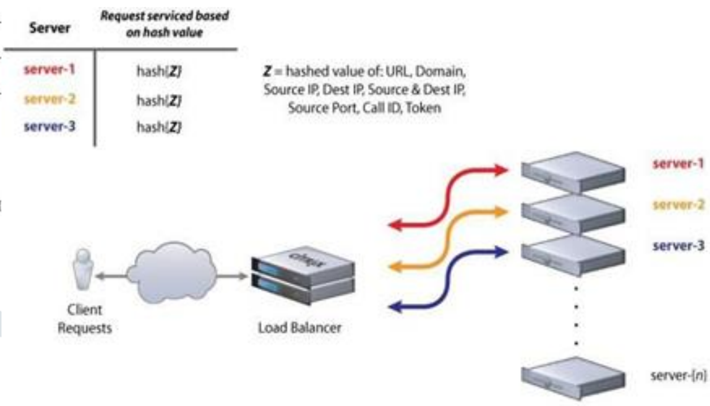


Figure 8: Source IP Hash [4]

2.10 Ant Colony:

This basic idea of ACO is to simulate the foraging behavior of ant colonies. When an ant's groups try to search for the food, they use a special kind of chemical to communicate with each other. That chemical is referred to as pheromone. Initially ants start search their foods randomly. Once the ants find a path to food source, they leave pheromone on the path. An ant can follow the trails of the other ants to the food source by sensing pheromone on the ground. As this process continues, most of the ants attract to choose the shortest path as there have been a huge amount of pheromones accumulated on this path. This collective pheromone depositing and pheromone. Following behavior of ants becomes the inspiring source of ACO [7].

ACO Algorithm approach: Initialization of algorithm: All pheromone values and parameters are initialized at the beginning of the algorithm [7]. Initialization of ants: M number of ants are initialized to select N number of tasks. Each ant build solution to M number of resources. In each iteration ants are randomly selected to build constructive direction [7]. Local Pheromone Updating: After M ants map solution to M number of resources pheromone value is updated by local pheromone updating rule. The local pheromone update is performed by all the ants after each construction step. Each ant applies it only to the last edge traversed [7].

Limitations of ACO [7]: Recruitment Strategies (Methods used to communicate previous search experiences to other members of the colony) are indirect. Exploration may not be sufficient It is more adaptive in Nature. Its convergence is guaranteed but time to convergence is uncertain. Coding is not straightforward. It is prone to falling in the local optimal solution.

2.11 Load Balance Improved Min-Min Scheduling Algorithm (LBIMM): It starts by executing Min-Min algorithm at the first step. At the second step it chooses the smallest size task fro0m the heaviest loaded resource and calculates the completion time for that task on all other resources. Then the minimum completion time of that the compared with the make span produced by Min-Min. This makes LBIMM to produce a schedule which improves load balancing and also reduces the overall completion time. But still it does not consider priority of a job while scheduling [8].

2.13 Particle Swarm Optimization: Particle Swarm Optimization algorithm is a new kind of modern heuristics algorithm and it is well characterized by its self-organizing, self-learning, self-adaptive characteristics and the implicit parallelism and guided search, which is often used to solve different kinds of NP-complete problems and complex task scheduling problems. The particle swarm optimization (PSO) is a population-based algorithm that was invented by Kennedy and Eberhart, which was inspired by the social behavior of animals such as fish schooling and bird flocking. Another advantage of PSO is that has very few parameters to adjust which makes it particularly easy to implement. It was pointed out that although PSO may outperform other evolutionary algorithms in the early iterations, its performance may not be competitive as the number generations are increased [9].
The PSO algorithm has shown its robustness and efficiency in solving function value optimization problems and many other research fields, such as practices of neural network, control of fuzzy system etc. In this paper, a new PSO with mutation operator (PSOMO) is presented based on the standard PSO. By introducing mutation operation, the algorithm shows higher calculation efficiency and a faster convergence rate than PSO [10].

3. RELATED WORK

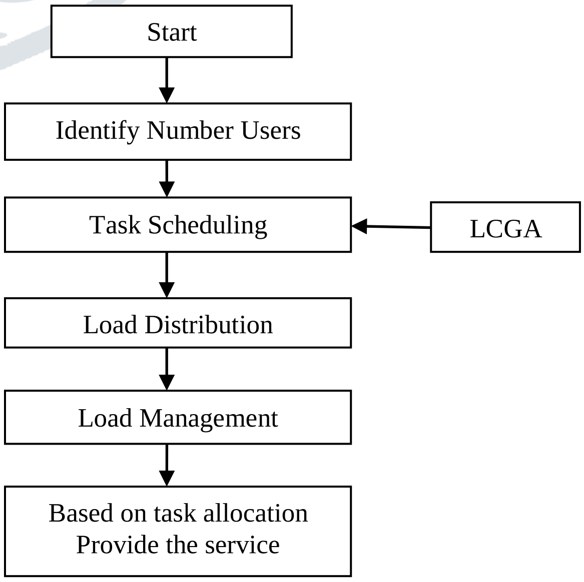
(Shuang YIN, Peng KE, Ling TAO, IEEE,2018 [3]), In this paper Efficient task scheduling mechanism not only meets users' requirements but also ensures cloud resources' high utilization. A new scheduling algorithm based on double-fitness algorithm-load balancing and task completion cost genetic algorithm (LCGA) is proposed. The scheduling guarantee load balancing and makes task completion cost less. At the same time, this paper brings in not just variance to represent the load among computing workers but weights multi-fitness function. (Manas M N, Nagalakshmi C K, Shobha G, ISSN: 2319-5940, April 2014 [1]), In this paper is load balancing basically balances the load to achieve higher throughput and better resource utilization. Since scheduling task is NP- complete problem, so heuristic and meta heuristic approaches are preferred options to solve the same. In this paper, they chose a meta- heuristic approach of Ant colony optimization algorithm to solve the task scheduling problem in cloud environment focusing mainly on two objectives, i.e., minimizing the make span/ computation time and better load balancing. (Sameeksha Irkal, Vandana Sunag IJIRCCE, ISSN (Print): 2320-9798,2017 [4]), Cloud computing is a technology which provides various computing services like a server, database, networking, software analytics and many such services over the internet on a Pay-per-use-on-demand model. With the major advancements and success of cloud, it is attracting more users and enterprises to store and process their data on the cloud thereby reducing the infrastructure costs and at the same time achieving data integrity, scalability, reliability, high productivity and security at minimal costs.

4. PROPOSED METHODOLOGY

This Section we are using for methodology of existing system and proposed system for dissertation mainly works of load balancing, Task scheduling of cloud computing. Basically, we are work in task scheduling, load distribution, load management and task allocation in cloud computing. We have drawn the flow of existing and proposed system and follow steps of algorithm and utilize methods.

4.1 Existing System:

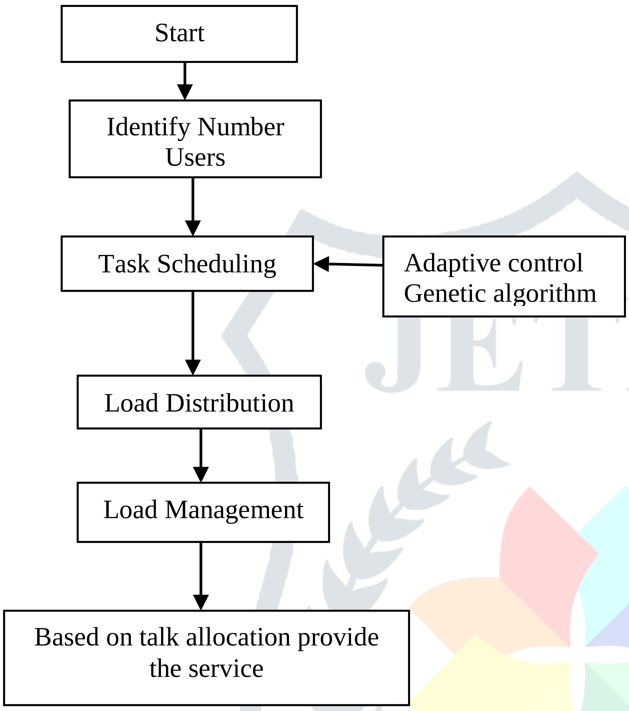
Algorithm:
Steps 1: User Task has been started like data, information, processes, resources etc.
Steps 2: Cloud computing infrastructure is composed of a large number of server and a single bare metal cloud system. One or more Users can access the required resources through the network; they don't know how to call this internal cloud computing infrastructure.



Step 3: Task Scheduling refers to the sequencing of how tasks have been scheduled for completion over a period of

time to ensure that the tasks get done in the most appropriate order based on priority and other considerations and that any inter-dependencies among tasks are taken into consideration. They extend GA and develop a new algorithm that is called Local Cultivation Genetic Algorithm (LCGA). LCGA has a neighborhood crossover mechanism in addition to the mechanisms of Gas.

4.2 Proposed System:



5. CONCLUSION

Cloud Computing is a huge idea and Data security and Load balancing plays a very important role. As we all know, there are abundant advantage with load balancing, task scheduling for IT Environments, especially the dynamic implementation. It designed to help organizations achieve their availability objectives. Security is one of the major issues, when outsourcing highly sensitive data to the cloud. Load balancing helps with complete failover capabilities in case of task computing issue, server failures, distribution of traffic across multiple server and disaster recovery.

6. REFERENCES

1. Manas M N, Nagalakshmi C K, Shobha G,"Cloud Computing Security Issues And Methods to Overcome", ISSN (Print) : 2319-5940, April 2014.
2. Santosh Kumar, R H Goudar, " Cloud Computing – Research Issues, Challenge, Architecture, Platform and Application: A Survey" IJFCC ,2012
3. Shuang YIN, Peng KE, Ling TAO," An improved genetic algorithm for task scheduling in cloud computing", IEEE,2018.
4. Sameeksha Irkal, Vandana Sunag," Load Balancing Algorithms in Cloud Computing", IJIRCCE, ISSN (Print): 2320-9798,2017.
5. Alok singh, Vikas Kumar Tiwari, Dr. Bhupesh Gour,"A Survey on Load Balancing in Cloud Computing Using Soft Computing Technique's", IJARCCE,ISSN(Print) : 2319-594, Vol.3, Issue 9,September 2014.
6. S.Jyothsna," Distributed Load Balancing in Cloud using Honey Bee Optimization", IJETICS, ISSN 2278-6856, 2016.
7. Geetha Megharaj, Dr. Mohan K.G," A Survey on Load Balancing Techniques in Cloud Computing", ISOR-JCE, ISSN: 2278-8727,2016
8. Tao Jiang*, Jia Li," Research on the Task Scheduling Algorithm for Cloud Computing on the Basis of Particle Swarm Optimization", IJSSST, 2011
9. Arjun U, Vinay S," A Short Review on Data Security and Privacy Issues in Cloud Computing", IEEE, 2016.
10. Mayanka Katyal*, Atul Mishra," A Comparative Study of Load Balancing Algorithms in Cloud Computing Environment".