# Comparative performance analysis for the load balancing mechanism in the Cloud computing environment using optimization methods

**Nand Kurmi**

*Abstract— Load balancing plays a very big role within the field of cloud computing, that ensures the efficient and truthful distribution of the resources among the distributed computing .Cloud computing provides us a computing model that relies on the web, the information square measure very vast and is present in large amount therefore. It must be forced to improve and maintain the policy of planning and exchange of resources efficiently. during this paper, a model has been proposed for the load leveling mechanism during a cloud computing atmosphere that is predicated on the swarm intelligence family ways like particle swarm optimization, that is compared with alternative techniques like round robin method, equally spread the current execution load and additionally With some transformative strategies like the genetic algorithm. Our empirical simulation and result analysis study shows that our proposed methodology produces higher results than other existing techniques. This work is on the basis of time for every rule and chooses the simplest ways for the several issues.*

*Index Terms— cloud computing, virtual machine, load balancing, Iaas, Paas, Saas, Resource utilization.*

## I. INTRODUCTION

Cloud computing offers a large number of helpful services to share a large amount of data, an outsized storage capacity, various IT resources and an in depth data of the analysis. Though applications got to recover information from distributed storage, the bandwidth between process nodes and Storage nodes will have a particular impact on application execution once the system status is shaky. The definition of the National Institute of Standards and Technology (NIST) states: "Cloud computing can be a model to allow access to the current, convenient. And free quickly with minimal service management or interaction effort from the service provider. "The load balancing mechanism is an area for transmitting requests over an outsized quantity of resources and helps the network avoid annoying period of time and provides best performance for users. It's the process of distributing the load between completely different nodes or devices of distributed systems that must improve the utilization of resources and therefore the interval of the work, avoiding even a state of affairs of filling a given node with an important load. These schemes make sure that all processors within the system or every device within the network perform just about associate degree equal quantity of work at any time. Cloud computing provides the services on demand using internet connection therefore, to efficiently share the resources among all devices in a distributed manner is a very challenging task.
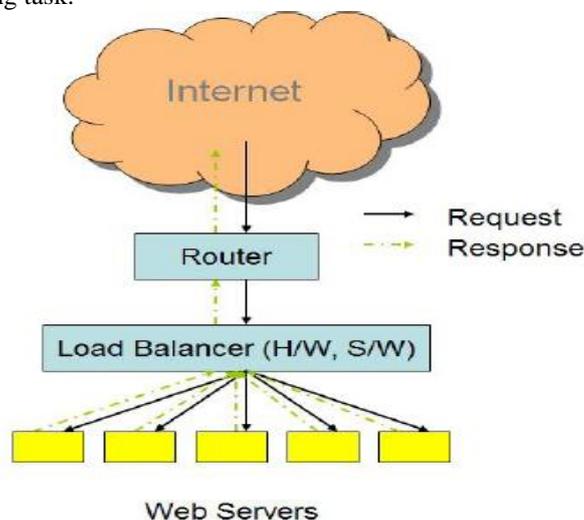


**Fig 1: Workflow of network load balancing**

The above diagram depicts the workflow of load balancing mechanism in which integrate all the hardware and software resources with using internet from the help of some connecting internetworking devices like router and nay other device. Load balancing is one of the techniques which provide the efficient management of resources, it is a technique to distribute workload among all servers and network interfaces. It also offers new opportunities and scales of economies as well as to present their unique set of challenges.
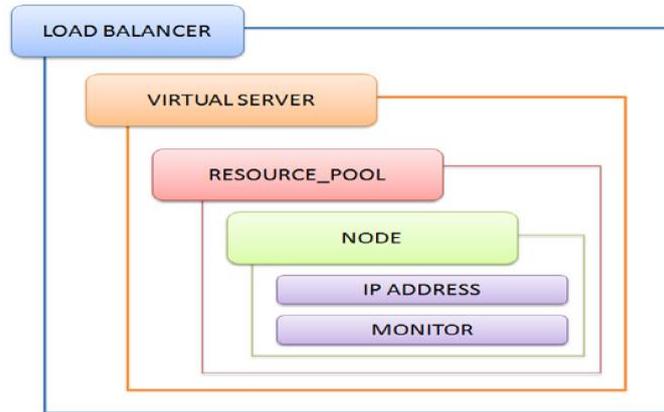


**Fig 2: Load Balancer**

Load balancing is the process of delivering loads across all servers, such that no server should be overwhelmed by jobs or tasks. A cloud data centre holds a set of servers (or hosts). Whenever a client makes a request to do a particular job, the request is directed to a server with sufficient resources to do that job. In return, the request is handled by a suitable virtual machine to run on it. Since the requests rose by the customers increase, the load on the server also increases.

*CLOUD COMPUTING BENEFITS*

**1) Scalability**–With these positive aspects of cloud computing, collectively, using cloud computing, there are several challenges in which the research industry offers many resources to provide scalability, load balancing, quality of quality (QoS), application performance. Cloud computing provides a better scalability, always available method for accessing data and resources.

**2) Improved reliability**– Cloud computing can cost-effective most of the unnecessary sites, making business stability easier.

**3) Less infrastructure costs**– Clients do not have their infrastructure with external clouds, this facilitates enterprise to remove capital expenditure and utilize resources in the form of services, which they use only. Cloud applications allow saving of implementation and security costs.

**4) Enhance utilization**– Cloud computing is able to increase usage rates by sharing data with the same computing power between multiple clients, in addition to reducing information technology basic cost.

**5) Improve end-user productivity**–Cloud computing provides, users can use any device, such as system or mobile, notebook, despite their location.

**6) Highly secure**– Security concerns are a compulsory issue in cloud computing, due to virtualization, which is common to all cloud models, important security threats for virtual data center and cloud infrastructure, VM theft, VM avoidance, hyper-jacking, data leakage and rejection services other concerns are abundance, velocity of attack, and data privacy etc.

**7) Energy efficient**– Cloud-based computing IT reduces capital costs, reduces labor costs, increases productivity and saves around 68-87% of energy, office computing reduces the same amount of carbon emissions.

## II. LITERATURE SURVEY

**Yasser Alharbi and Kun Yang (2016)[1]** proposed the virtual machine game plan that considers both figuring resources and I/O data. The point of this calculation is to decrease the general occupation finishing time, which incorporates figuring time and also information exchanging them. They present here the asset calculation and sharing instrument for the proficient load adjusting.

**Amir Nahir, Ariel Orda and Danny Raz(2016) [2]**Analyze the present scheduler plot which frequently takes high correspondence overhead when gathering the information expected to settle on booking choices, so there are delays in employment asks for on their way to the execution server. In this way, they offer a novel arrangement

that does not overhead any correspondence among clients and servers on the landing of an occupation, in this manner killing any programming over-burden on the imperative work way. The methodology depends on the conviction of different generations of each activity and furthermore the causation of each reproduction on a one of a kind server.

**Matthew Malensek, Sangmi Pallickara, and Shrideep Pallickara (2016) [3]** Cloud has turned into a moderate and adaptable answer for foundation necessities, enabling you to combine a wide range of figuring assets into a physical machine. Notwithstanding, the arrangement of multi-center CPUs, memory modules with high ability and virtualization innovations have amplified the thickness of the incorporated framework, the circle I/O has turned into an awesome obstruction, particularly with regards to high limit mechanical plates Here, the creators analyze that in the information serious registering condition, bolstered by prescient model programming, anticipated models, and customer side coordination, the group can influence generally speaking execution and reaction. They assess their structure with a guide that decreases delegate work in the group of 1,200 virtual machines, which demonstrates a 21% change in the fruition of overwhelming circle debate.

**Alaka Ananth and K. Chandrasekaran (2015)[4]** proposed the strategy when numerous clients ask for administrations, the cloud specialist organization needs to legitimately set up solicitations for proper assets to meet every client's demand keeping in mind the end goal to meet administration level assertions (SLA) and hindrances as far as possible is. Occupation arranging is a critical issue in the cloud condition, where the principle objective is to decide the work for an effective utilization of assets. To get the most extreme advantage from the administrations gave to the shoppers. The cloud specialist co-op has the most ideal booking. The point of this paper is to augment the benefit of the specialist organization by boosting asset utilization and in addition the helpful diversion hypothesis based methodology for occupation planning for the cloud condition. What's more, it additionally centers on time restrict infringement and lessening work for the client submitted occupations.

**Saurabh Kumar Garg, Adel Nadjaran Toosi, Srinivasa K. Gopalaiyengar and Rajkumar Buyya(2014)[5]** The examination work incorporated the headway of a capable VM stack adjusting estimation for the cloud and the execution of a close examination of the proposed computation with the present figurings. Middle of the road expectations included concentrate the current VM stack adjusting calculations, proposing a proficient calculation for VM stack adjusting, actualizing the calculation on Cloud Analyst, and contrasting the proposed calculation and the current calculations on distinguishing parameters.

**Abhishek Gupta, Laxmikant V. Kale, Dejan Milojicic, Paolo Faraboschi and Susanne M. Balle(2013)[6]** In this examination, they address the cognizant designation of the application from n VM occurrences (counting a solitary employment ask for) by outlining and actualizing a Programmer good with HPC in OpenStack Compute (NOVA) and joining it in the test system (cloud System). Through different enhancements, specifically the learning of topology and equipment, the bookkeeping of impedances between virtual machines and the cognizant combination of the application, show a progressed VM situating, with a 45% change in HPC execution and/or 32% Limit the expansion of the flimsiness (or commotion) impact up to 8%.

**Gao, Yanzhi Wang, Sandeep K. Gupta and Massoud Pedram(2013)[7]** they has recommended that customers can make their very own organizations and applications in perspective of open virtual sets Machines are liberated from the weight of assets Provisioning and Work Scheduling Then there will be gaining by CSP Each client oversees information parallelism in a remaining task at hand viably To apply aggregate client solicitations, and custom customization Create worldwide vitality expenses and due dates mindful cloud stages.

**Zeratul Izzah Mohd Yusoh and Maolin Tang (2012) [8]** as an administration programming (saas) programming is accepting greatest consideration from clients and suppliers. It has raised numerous new difficulties by giving better SaaS in SaaS providers, which is fundamental for all with negligible expenses. One of the rising ways to deal with managing the test is to disperse SaaS. In such a methodology, there are numerous points of interest, including the adaptable offer of SaaS capacities and a decrease in the expense of enrollment for clients, to appropriate it. Notwithstanding, this methodology additionally shows new issues for overseeing SaaS assets in the cloud server farm. here present the issue of in general SaaS assets administration in the cloud server farm, particularly on its underlying arrangement and asset improvement issues, with the point of decreasing asset utilization, and in addition enhancing SaaS execution based on execution time.

**Kien Le, Jingru Zhang, Jiandong Meng, Ricardo Bianchini ,Yogesh Jaluria and Thu D. Nguyen (2011)[9]**Cloud pro associations administer various geographically scattered information server . These server ranches consume a ton of vitality, which changes over into high working expense. The fascinating thing is that the land appropriation of server farms offers numerous open doors for reserve funds. Here, they consider the impact of cooling load arranging systems and the most extraordinary server farm temperature. In context of this examination suggestion and on the dynamic load dispersing plans considered as a transient cooling effect and all related power costs.

**Kyong Hoon Kim, Anton Beloglazov and Rajkumar Buyya (2011)[10]** the paper writer showed the arrangement and use of the outline and execution structure for distributed computing administrations organizations. Your framework adaptively adjusts virtual physical assets in light of the changing needs of different sites. The framework oversees both to abstain from over-burdening and green registering for frameworks with various asset imperatives. As a future work on this venture, we add greater security to our current work and furthermore to incorporate the portable condition where ideal utilization of assets is required.

**George Kousiouris, Tommaso Cucinotta and Theodora Varvarigou(2011) [11]** The goal of this report is to ponder and foresee the impact of some vital parameters in the execution of virtual machines (VMs).These parameters incorporate distribution rates, ongoing booking choices, and VM co-positions, when they are sent on the equivalent physical hub, as controlled by the server union pattern and late advancement in the distributed computing framework is finished. Different mixes of VM outstanding task at hand composes are analyzed in connection to the previously mentioned variables to discover ideal allotment methodologies. At last, a black box in view of hereditarily upgraded counterfeit neural systems is embedded to check the absence of prioritization of execution and it is contrasted and straight relapse technique.

**Sriram Kailasam, Nathan Gnanasambandam, Janakiram Dharanipragada and Naveen Sharma(2010)[12]** The act of processing in at least two server farms isolated by the Internet is expanding in notoriety because of adaptable registering requests and blasts in the compensation as-you-go plans given on the cloud. Cloud is tending to this procedure of scaling all over in the server farms offering ensured benefit level, particularly the best exertion movement and a major test for expansive cloud-based distributed computing. Parallel outstanding burden that has been addresses ongoing and incorporate between cloud handling and examination of picture records. To determine the issues, this examination offer three kinds of Autonomous Cloud-Fluent Scheduler, which will enhance the remaining burden attributes, data transfer capacity and accessibility of accessible assets in the administration levels required by clients, (for example, accelerate and line succession insurance) To give assets accessible. Specifically, these go getter schedulers utilize time-arrangement grouping for archive preparing remaining tasks at hand.

**Jing Tai Piao and Jun Yan (2010) [13]** this regularly impacts the general execution of the application with the I/O execution of the information arrange. Along these lines, the assignment of virtual machines that host applications and move to them. Virtual machines, regardless of surprising system inactivity or blockage, are imperative for keeping up application execution. To take care of these issues, they propose a virtual machine. Confinement and relocation ways to deal with lessen information exchange times.

**Po-Huei Liang and Jiann-Min Yang[14]** This report introduces a structure for worldwide load leveling of the site server amid a cloud with two or level load leveling model. Accessibility and dependability of the customer Systems will dependably be more imperative. These days the cloud applies the littlest interference, the pixie demonstration It will be huge for clients. It is an essential issue that ensures dependability and soundness. from the cloud Load adjusting locales It could be a decent arrangement. The proposed structure is intended to adjust to heap of inception shaking framework and system the activity permits the system specialist co-op to execute a heap adjusting administration in various information it concentrates progressively, while clients require more load balancers to expand accessibility.

### III.  PROPOSED METHOD

The currently available algorithms do not save the state of the previous virtual machine assignment on a request from a particular user base. Therefore, each time a request is received from the same user base, the algorithm must be executed again, which increases the total response time and request processing time. This research work aims

to reduce the total response time and processing time of basic user requests by proposing particle swarm optimization with the server affinity VM load balancing algorithm.
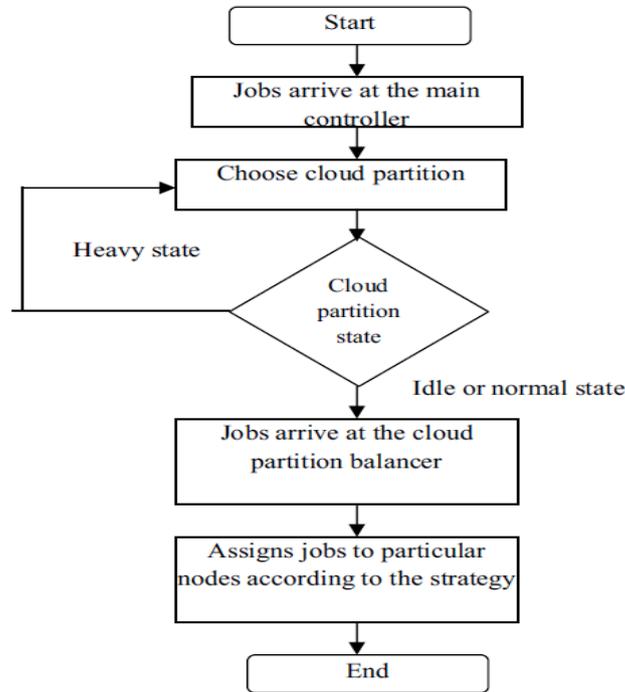


Fig 3: Job allocation processes in public cloud computing model.

1. Generate n solutions for random population.
2. For each individual x € N: calculate fitness of each particle (x).
3. Initialize value of weight factor, w.
4. For x = 1 on n particles

   Set pBest as the best position of the particle x

   If fitness (x) is better than pBest then

   Set pBest (x) = fitness (x)

   End for x
5. Establish gBest as the best physical form of all particles
6. For x = from 1 to n particles
7. Update the weight value factor w
8. Check if the termination is true.
9. End

*COUNTING PRIMITIVE OPERATIONS*

The numbers of operations are performed in above algorithm as:

Step 1 contributes one operation for n times.

Step 2 contributes one operation for n times.

Step 3 requires one operation to count for initialization

Step 4 performs n iteration in outer loop and using bubble sort for inner loop is n.log (n).

Step 5 finds best in n particles in log (n) operations.

Step 6 performs n loop with 14 operations.

Step 7 updates weight factor in one operation.

$T (n) = n+n+1+n. \log (n) + \log (n) \ 14+1.$

$T (n) = 2n+16+ (n.\log (n).t).$

## IV.   EXPERIMENTAL RESULT ANALYSIS

A cloud segment is a sub-area of the general cloud of the population with divisions in sight of the geographical areas. The load distribution method relies upon the concept of isolating the load. Amid this paper, we have a

tendency to arrange a relative examination model of load adjusting in distributed computing. There are different kinds of procedures, for example, Round Robin, Equality space current execution load and molecule of swarm optimization etc.We compare all the techniques and find that the particle of swarm optimization better results than other existing techniques as a proposed strategy.
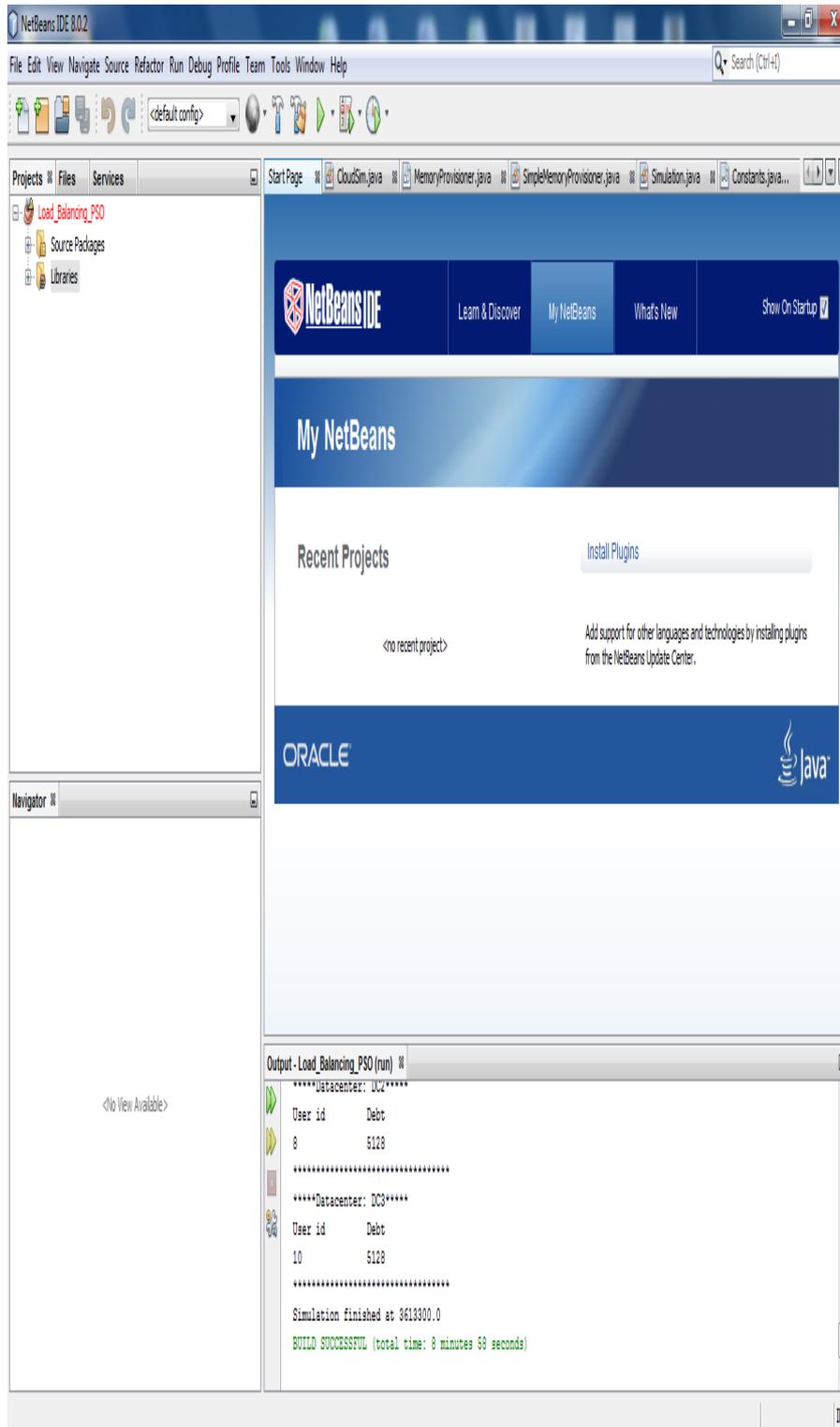


**Fig 4: Implementation window for the experimental process using in a load balancing.**
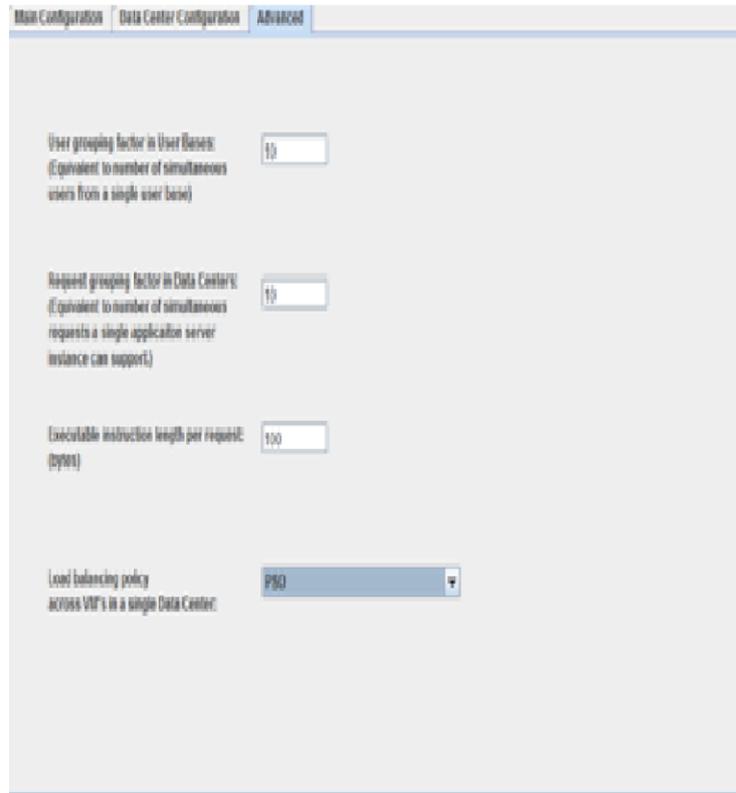
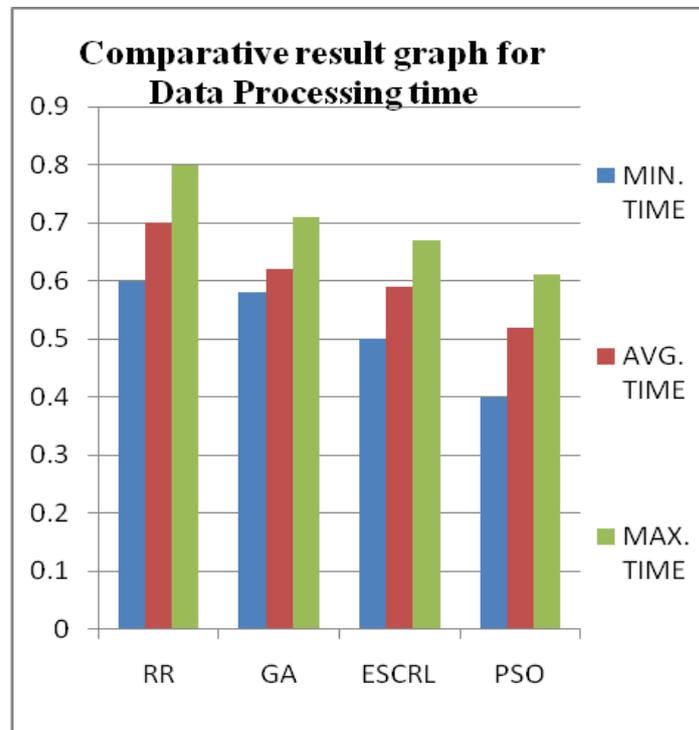**Fig 5: Implementation window with method using in a load balancing.**



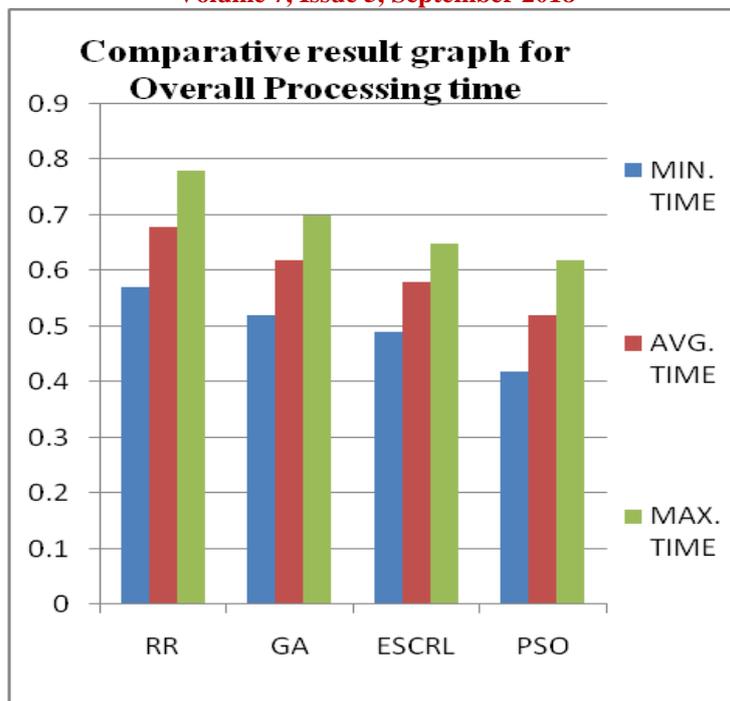**Fig 6: Comparative Result graph for the Data processing time with methods.**

**Fig 7: Comparative Result graph for the overall processing time with methods**

## V. CONCLUSION AND FUTURE WORK

Cloud computing plays a very important role in the modern age of computing, since it provides a large number of user-level services on demand, the number of active users can try to try the services simultaneously. In this article, we highlight the comparative study of different load balancing algorithms to find the use of the user base and the data center for each algorithm. In future work, we use a more meta-heuristic algorithm to improve the experimental results in terms of calculation time.

## REFERENCES

[1] Yasser Alharbi and Kun Yang "Optimizing jobs completion time in cloud systems during Virtual Machine Placement", International Conference on Big Data and Smart City, 2016, Pp 1-6.

[2] Amir Nahir, Ariel Orda and Danny Raz "Replication-Based Load Balancing", IEEE, 2016, Pp 494-507.

[3] Matthew Malensek, Sangmi Pallickara, and Shrideep Pallickara "Minerva: Proactive Disk Scheduling for QoS in Multitier, Multitenant Cloud Environments", IEEE, 2016, Pp 19-27.

[4] Alaka Ananth and K. Chandrasekaran "Cooperative Game Theoretic Approach for Job Scheduling in Cloud Computing", Computing and Network Communications, 2015, Pp 147-156.

[5] Saurabh Kumar Garg, Adel Nadjaran Toosi, Srinivasa K. Gopalaiyengar and Rajkumar Buyya "SLA-based virtual machine management for heterogeneous workloads in a cloud datacenter", Journal of Network and Computer Applications, 2014, Pp 108-119.

[6] Abhishek Gupta, Laxmikant V. Kale, Dejan Milojicic, Paolo Faraboschi and Susanne M. Balle "HPC-Aware VM Placement in Infrastructure Clouds", IEEE, 2013, Pp 11-20.

[7] Yue Gao, Yanzhi Wang, Sandeep K. Gupta and Massoud Pedram "An Energy and Deadline Aware Resource Provisioning, Scheduling and Optimization Framework for Cloud Systems", IEEE, 2013, Pp 1-10.

[8] Zeratul Izzah Mohd Yusoh and Maolin Tang "Composite SaaS Placement and Resource Optimization in Cloud Computing using Evolutionary Algorithms", IEEE, 2012, Pp 590-597.

[9] Kien Le, Jingru Zhang, Jiandong Meng, Ricardo Bianchini, Yogesh Jaluria and Thu D. Nguyen "Reducing Electricity Cost Through Virtual Machine Placement in High Performance Computing Clouds", ACM, 2011, Pp 1-12.

[10] Kyong Hoon Kim, Anton Beloglazov and Rajkumar Buyya "Power-Aware Provisioning of Virtual Machines for Real-Time Cloud Services", John Wiley & Sons, Ltd., 2011, Pp 1-19.

[11] George Kousiouris, Tommaso Cucinotta and Theodora Varvarigou "The effects of scheduling, workload type and consolidation scenarios on virtual machine performance and their prediction through optimized artificial neural networks", The Journal of Systems and Software, 2011, Pp 1270-1291.

[12] Sriram Kailasam, Nathan Gnanasambandam, Janakiram Dharanipragada and Naveen Sharma "Optimizing Service Level Agreements for Autonomic Cloud Bursting Schedulers", Parallel Processing Workshops, 2010, Pp 285-294.

[13] Jing Tai Piao and Jun Yan "A Network-aware Virtual Machine Placement and Migration Approach in Cloud Computing", IEEE, 2010, Pp 87-92.

[14] Po-Huei Liang and Jiann-Min Yang "Evaluation Of Two-Level Global Load Balancing Framework In Cloud Environment" International Journal of Computer Science & Information Technology (IJCSIT) Vol 7, No 2, April 2015. Pp 1-11.