Volume 5, Issue 8, August 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Load Balancing in Cloud Computing

Durairaj.M
Assistant Professor, School of Computer Science, Engineering and Applications, Bharathidasan University, Trichy, India

Menaka.A
Research Scholar, School of Computer Science, Engineering and Applications, Bharathidasan University, Trichy, India

Abstract— Load balancing in cloud computing is emerging area of research. Cloud user gets better balancing services through better load balancing which resulted in performance improvement in cloud. Load balancing operation will be initiated to effectively use the resources dynamically with the process of assigning resources to the corresponding node to reduce the load values. Load balancing is a heart of every computer model, with load balancing of some kind you can't scale and scalability is one benefits of cloud. Cloud computing is known for its on demand service as it offers efficient resource allocation and also offers dynamic, flexible and reliable services to the customers through pay-as-you use method. This paper presents a review of load balancing algorithms and techniques applied in cloud computing, which include dynamic load balancing strategies directed on cloud data storage based on workload. This work also aims to present execution analysis of various load balancing algorithms.

Keywords— Cloud computing, Load balancing, Static and Dynamic Algorithm, Simulation

I. INTRODUCTION

Cloud computing is an essential ingredient of advanced computing system. Anywhere user is able to rent software and hardware infrastructure and computational resources as per user basic computing concept, technology, architectures that have developed and consolidated in the last decades. Anywhere in the world, Cloud computing allow us to access all our application and documents, freeing from the confines of the desktop and making it easier for group members in different locations to collaborate. Cloud computing is not a network computing. Balancing in cloud computing provides an efficient solution to various issues residing in cloud computing environment set-up and usage.

In Cloud computing, multiple cloud users can request number of cloud services simultaneously. All resources in cloud must be a provision that is made available to requesting user in efficient manner to satisfy their need without compromising on the performance of the resources. This model is scalable enough to represent systems composed of thousands of resources and it makes possible to represent both physical and virtual resources exploiting cloud specific concepts such as the infrastructure elasticity. Load balancing is a methodology to distribute workload across multiple computers or other resources over the network links to achieve optimal resource utilization, maximize and minimum response times, and avoid overload.

A. LOAD BALANCING

Load Balancing is essential for efficient operations in distributed environments. As Cloud Computing is the greatest platform which provides storage of data in very lower cost and available for all time over the internet, that's why load balancing for the cloud has become a very interesting and important research area. Load balancing helps to attain a high user satisfaction and resource utilization ratio by ensuring an efficient and fair allocation of every computing resource. Load balancing of course, distributes load across multiple instance of an application to improve the performance, and maintains availability which enables the scale. In most cloud environments, where provider supplied load balancing service, are based on a scale out model, means that scalability is based purely on doing of new application instance when demand reaches a certain threshold [4].

B. LOAD BALANCER

The Load Balancer systems allow you to create an infrastructure able to distribute the balancing between two or more Cloud Servers. You can therefore shape your infrastructure to allow sustaining activity peaks, optimizing the allocation of resources and ensuring a minimal response time. Using a load balancer is recommended in all cases whether you require one or more of the following:

- High traffic and request peaks.
- Guarantee of service continuities.
- Specialization of servers.

This paper focuses on reviewing some of the existing cloud computing works found on the recently published works on reducing load balancing and estimating cost. This paper organized as the section 1 gives introduction, and section 2 briefs the techniques and service models of cloud computing. Section 3 presents the architecture of a service also

discusses different service models. Section 4 presents brief review of existing works on recently published papers. Section 5 discusses and cons of the reviewed works. Section 6 concludes with the future direction of this work.

II. RELATED WORK

Mezmaza, *et al.*, 2013[1] presented a decentralized architecture of the energy aware resource management system for cloud data canters while meeting QoS requirements. There researchers present heuristics and three stages of continuous optimization of VM (Virtual Machine) placement. Heuristics have been evaluated by simulation using extended CloudSim toolkit in heterogeneous workload independent environment of virtual machines (VMs).

Hung, *et al*., 2013[2] elaborated a dynamic management is an active area of research. The authors presented prediction techniques and queuing theory results to allocate resources efficiently within a single server serving a web workload. The static allocation approaches used in where authors propose a simple heuristic for vector been packing problem and apply it to minimize the number of servers required to host a given web traffic. The performance control of web server, the control theory is applied to design a system. Based on the feedback system, the arrival rate of requests to the server is throttled. Then the authors proposed an optimization algorithm that allocates resources (i.e., web servers) depending on the expected financial gain for the hosting canter.

Sridhar, *et al.*, 2010[3] discusses about the cost includes computational resource utilization and transferring binary and data the running instances of mobile application. This paper also presented an offloading framework for simple dynamic application profiling and Partitioning technique. Which increases the memory allocation, energy consumption of the mobile device is turnaround the time of the application. It is shown that application profiling requires an additional 8192KB RAM for maintaining at temporary trace file.

Soto mayor, *et al.*, 2009[4] discussed the presented load balancing techniques in cloud computing in their paper. Two main types of load balancing algorithms are Static load balancing and Dynamic load balancing. A static well-known load balancing technique called Round Robin, in which all processes are divided amid all available processors. The allocation order of processes is maintained locally which is independent of the allocation from the remote processor.

Radojevic, *et al*., 2013[5] focused in their work on variability and evolution over time. For accessing the performance of just focusing on the (average) performance values. This work characterizes the long-term performance variability of production cloud services close to this work is the seminal study of Amazon S3, which also includes a 40 days evaluation of the service availability. This work complements study by analysing the performance of eight other AWS and GAE services over a year.

Wang, *et al.*, 2010[6] presented a dynamic balancing algorithm called load balancing Min-Min (LBMM) technique, which is based on three level frameworks. This technique uses opportunistic load balancing algorithm, which keep each node busy in the cloud without considering execution time of node. Because of this it causes bottle neck in system.

Hamada, *et al.*, 2012[7] clarified in their paper that the cloud computing cannot be considered as a or technology that arose in recent years; instead, its root can be found in what John McCarthy described as the ability to provide computational resources as a utility Based on standard material presented by NIST (National Institute of Standards and Technology). Cloud computing is composed of five main characteristics, and two other characteristics are added based on the literature, with three spanning service models and four models of deployment.

Shiraz, *et al.*, 2014[8] discussed a mechanism of outsource. Computationally intensive applications entirely or partially to a remote server is called computational offloading. Ian MCC application, offloading is deployed for coping with the challenge of executing on SMDs and computationally intensive applications. A number of computational offloading frameworks have been proposed in the recent years, for computationally intensive mobile applications, which are elastic in nature. Elastic applications are partitioned at different granularity level strum time for the establishment of a distributed processing platform.

Indrajit Roy, *et al.*, 2010[9] presented a system for distributed computations using a combination of mandatory access control and differential privacy which provides end-to-end integrity confidentiality, and privacy guarantees. A Differential privacy is a methodology for ensuring the output of aggregate computations does not violate the privacy of individual inputs. It mathematically declassifies data in a mandatory access control system. Air vat enables the execution of trusted and UN trusted Map Reduce.

Huany, *et al.*, 2012[10] proposed a technique called dynamic self ballooning, where a driver the guest virtual machines, continuously reclaiming free and unused memory and giving it back to the hypervisor. This technique during migration reduces the amount of memory that is to be sent of the network since the reclaimed and unused memory is not transferred. Continues with pull phase the post-copy approach begins with the stop and copy phase. For the destination machine to have all the data it requires, it retrieves them from the source continually. Several techniques can be used in order to retrieve the memory from the source, such as demand paging, active pushing, and adaptive pre-paging. The migration time of post-copy is mostly bounded by the amount of memory allocated to the virtual machines since the memory is the bottleneck of saving and transferring the state. As opposed to pre-copy, post-copy will only transfer each memory page once.

Ellinger, *et al.*, 2013[11] developed control methods include providing resource guarantees for VMs in the form of reservations, enforcing resource limits with limits or maximums and manipulating dynamic resource scheduling priorities with shares are demonstrated. With reservations, the service providers or end users can explicitly specify the resources that are reserved for the deployed VMs. These reservations guarantee that the VM is entitled to the specified resources, and will receive at least the specified amount as long as it demands it. Resource guarantees are commonly employed to

impose service levels, similar to our performance-based capacity provisioning method. Moreover the described control mechanisms can also be dynamically modulated and applied to runtime, autonomic management.

Escalate, *et al.,* 2013[12] presented of research in energy efficient proposed architectural principles for energy efficient management of clouds, energy efficient VMs allocation policies and scheduling algorithms considering QoS expectations and power usage characteristics of the devices. It is validated by conducting a performance evaluation study using CloudSim toolkit.

III. METHODOLOGY

A. ROUND ROBIN ALGORITHM

It is one of the simplest scheduling techniques that utilize the principle of time slices. The time is divided into multiple slices and each node is given a particular time slice or instance interval i.e. it utilizes the rule of time scheduling. Every node is known a quantum with its operation. The resources of the repair provider are providing to the requesting client on the basis of time slice [22].

Step 1: Work arrives on the main controller.
Step 2: Work assigns near balancer according to balancer status, request location.
Step 3: In particular partition P
Set i=0 set s[n] as number of servers arranged in increasing order of jobs in P used for all n=1,.., n
Step 4: Set s[c] as number of idle servers in P from s[n] used for all c=1,.. , n
Step 5: When job arrives If s[c]! = NULL then, send the connection to s[i]
Step 6: i= i+1;
Step 7: If i == c
Step 8: Then i=1
Step 9: Else go to game theory
Step 10: End If
Step 11: Repeat step 1

B. EQUALLY SPREAD CURRENT EXECUTION ALGORITHM

Equally Spread Current Execution (ESCE) algorithm is showing an improvement in response time also the processing moment in time. The job be evenly spread, the complete computing system is load balanced and no virtual machines are underutilized. Due to this benefit, present be a decrease in the virtual machine cost and the data transfer cost.

Step 1: Find the next available VM.
Step 2: Check for all current allocation count is less than max Length of VM list allocate.
Step 3: If available VM is not allocated create a new one.
Step 4: Count the active load on each VM.
Step 5: Return the id of those having least load.

IV. TOOL USED IN CLOUD

Firstly, selection of good tool is critically important for simulating large scale application, so apparently or researchers choose a tool that has easy to use. So this tool provides by Graphical User Interface (GUI) which comes with a Tool Kit by setting in various cloud parameter. The output provides by this tool in graphical representation which can be easily examined by researcher. Some of the feature of this tool is:

- Easy set up.
- Flexibility in Configuring the Cloud Analyst environment.
- Output is in graphical form. Response and data canter processing time function are performance evaluation parameters. The result after simulation helps a lot in improving quality of service.

A. MAIN COMPONENTS OF CLOUD ANALYST

- Region
- User base
- Internet
- Internet Cloudlet
- Data Canter Controller
- VM Load Balancer
- Cloud Application Broker

V. EXPERIMENTAL RESULTS

The proposed algorithm implemented for simulation. Java language is used for implementing load balancing algorithms. VM Load Balancing algorithm is applied for measuring overall response time of the cloud. In this minimum and maximum (ms) time different number of virtual machines are combined and analysed for evaluating performance. The model is simulated to analyse the response time and data centres processing time of proposed work [23].

Table 1: Closest Data Center

Overall Response Time Summary	*Average(ms)*	*Minimum(ms)*	*Maximum(ms)*
Overall Response Time	*318.48*	*150.11*	*630.11*
Data Center Processing	*0.26*	*0.02*	*0.65*

The overall response time and Data Center Processing time are tabulated under average, minimum and maximum policies as shown in Table 1. The overall response time of the closest data center average time is 318.48 ms, minimum time is 150.11ms and maximum time of overall response is 630.11ms. Data center processing average time is 0.26 ms, minimum time is 0.02 ms and the maximum time of data center processing time is 0.65 ms. The pattern of response time distribution has not changed much with the main difference being the number of smaller peaks being reduced as not all the traffic is directed at the same data center this time.

Table 2: Response Time by Region

User Base	***A**verage(ms)*	*Minimum(ms)*	*Maximum(ms)*
UB1	*199.646*	*150.113*	*239.116*
UB2	*300.837*	*238.67*	*378.159*
UB3	*499.692*	*397.618*	*630.114*

Table 2 shows the table representation of the user base UB1, UB2, and UB3 are under the policies of average (ms), minimum (ms), maximum (ms) response time by region. On User Base 1(UB1), average response time is 199.646 ms, minimum time response is 150.113 ms, and maximum time response is 239.116 ms. Similarly in the User Base 2(UB2) average time response is 300.837 ms, minimum time response 238.67 ms and also maximum time response 378.159 ms. User Base 3(UB3) has the average time response 499.692 ms, minimum time response is 397.618 ms and maximum time response is 630.114 ms. The user base 1(UB1) is the best time response in the time response region. The response times experienced by each user base are depicted graphically as shown in figure 1.

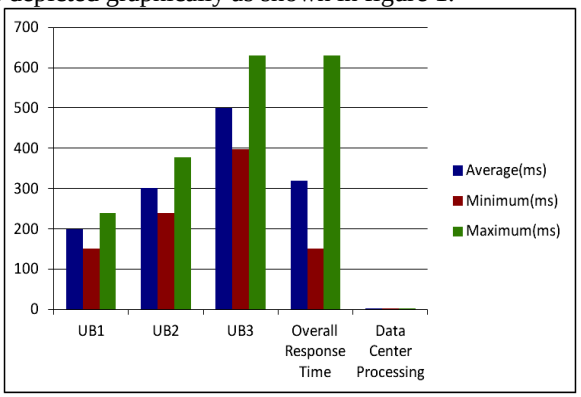

Fig 1: User Base and Average Response Times

Fig 1 in the graphical representation shows the average (ms), minimum (ms) and maximum (ms) for closest data center of three policies. UB1 average time is close to 200 ms, minimum time response is close to 150 ms and maximum time response is 230 ms. UB2 average response time is close to 300 ms, minimum time response is nearly 250 ms and maximum time response is close to 380 ms. Similarly, UB3 average response time is 500 ms, minimum response time is 400ms and maximum time 620 ms. Overall response time average is 320 ms, minimum time is 140 ms, and maximum time response is 620 ms. The data center processing average, minimum and maximum time response is under 1ms. The data center response time very is efficient compare to other time response.

Table 3: Optimize Response Time

Overall Response Time Summary	*Average(ms)*	*Minimum(ms)*	*Maximum(ms)*
Overall Response Time	*184.62*	*35.86*	*391.55*
Data Center Processing	*0.33*	*0.00*	*1.10*

The overall response time and Data Center Processing time are tabulated under average, minimum and maximum policies as shown in Table 3. In an optimize response time, the average time is 184.62 ms, minimum optimize response time is 35.86 and maximum optimize response time is 391.55 ms. In the data center processing optimizes response time, the average time is 0.33 ms and minimum time is 0.00 ms and maximum time is limit is 1.10 ms.

Table 4: Response Time by Region

User Base	*Average(ms)*	*Minimum(ms)*	*Maximum(ms)*
UB1	*50.128*	*35.856*	*65.358*
UB2	*200.101*	*146.116*	*267.111*
UB3	*300.077*	*208.593*	*391.546*

Table 4 shows the table representation of the user base UB1, UB2, and UB3 are under the policies of average (ms), minimum (ms), maximum (ms) response time by region. UB1 average time by region time is 50.123 ms, minimum time response region is 35.856 ms and maximum time is 65.358. UB2 response time region average time is 200.101 ms, minimum time is 146.116 ms and maximum time response is 267.111 ms. Similarly, UB3 response time region average time is 300.077 ms, minimum response time region is 208.593 ms and maximum response time region is 391.546 ms.

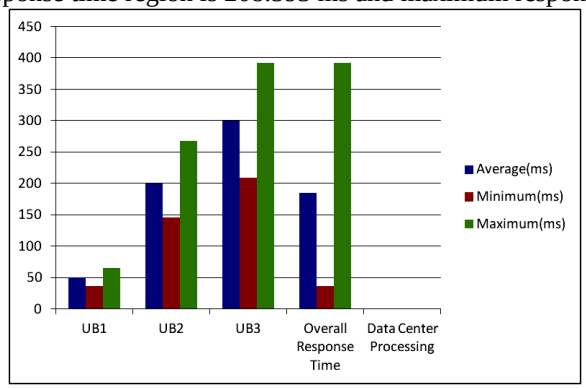

Fig 2: User Base and Average Response Times

Fig 2 shows the graphical representation of the average (ms), minimum (ms) and maximum (ms) of optimize response time of three policies. The graph shows that the UB1 average time is close to 50 ms, minimum time response is close to 35 ms and maximum time response is 70 ms average response time close is to 200 ms, minimum time response is nearly 140 ms and maximum time response is close to 270 ms. Similarly in UB3, the average response time is 300 ms, minimum response time is 200 ms and maximum time 380 ms. Overall response time average is 180 ms, minimum time is below 30 ms and maximum time response is 380 ms. The data center processing average, minimum and maximum time response is above 0.56 ms. The data center response time very efficient compare to other user base and response time.

Table 5: Reconfigure Dynamically With Load Balancing

Overall Response Time Summary	*Average(ms)*	*Minimum(ms)*	*Maximum(ms)*
Overall Response Time	*266.34*	*150.25*	*625.11*
Data Center Processing	*0.65*	*0.02*	*11.01*

The overall response time and Data Center Processing time are tabulated under average, minimum and maximum policies in Table 5. The reconfigure dynamically with load balancing average time is 266.34 ms, minimum time is 150.25 ms and maximum time of overall response is 625.11 ms. Data center processing average time is 0.65 ms, minimum time is 0.02 ms and the maximum time of data center processing time 11.01 ms.

Table 6: Response Time by Region

User Base	*Average(ms)*	*Minimum(ms)*	*Maximum(ms)*
UB1	*299.432*	*241.623*	*360.496*
UB2	*200.701*	*150.246*	*246.328*
UB3	*500.98*	*372.937*	*625.111*

Table 6 shows the table representation of the user base UB1, UB2, and UB3 are under the policies of average (ms), minimum (ms), maximum (ms) response time by region. On User Base 1(UB1), the average response time is 299.432 ms, minimum time response is 241.623 ms and maximum time response is 360.496 ms. Similarly, in the User Base 2 (UB2), average time response is 200.701 ms, minimum time response is 150.246 ms and also maximum time response is 372.937 ms and maximum time response is 625.111 ms. The User Base 1(ub1) is the best time response in the time response region. Region-wise response time distribution does not demonstrate any significant differences in the previous case, but the data center loading patterns show significant changes.

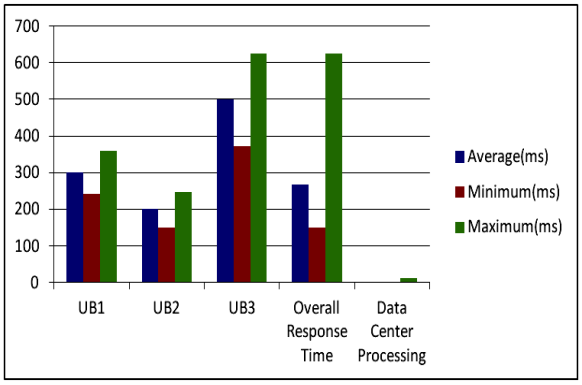

Fig 3: User Base and Average Response Times

Fig 3 shows the graphical representation of the average (ms), minimum (ms) and maximum (ms) for reconfigure dynamically with load balancing of three policies. The graph shows that the UB1 average time is close to 300 ms, minimum time response is close to 235 ms and maximum time response is 350 ms. UB2 average time is close to 200 ms, minimum time response is nearly 150 ms and maximum time response is close to 220 ms. Similarly UB3, the average response time is 500 ms, minimum time is below 150 ms and maximum time 620 ms. Overall response time average is 240 ms, minimum time is below 150 ms and maximum time response is 620 ms. The data center processing average, minimum and maximum time response is above 0.56 ms. The data center response time very efficient compare to other time response.

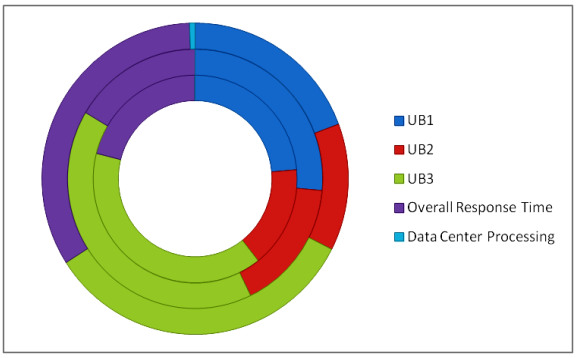

Fig 4 Comparison of Closest Data Center, Optimize Response Time and Reconfigure Dynamic with Load Balancing

The comparison of closest data center, optimize response time and Reconfigure dynamic with load balancing is illustrated in Fig 4.

VI. EXPERIMENTAL RESULTS

The problem of resource selection in distribute environment has received much attention. Here many of the previous mechanism, resource collection refer to the selection of computational resource in grid environment. The objective of this work is to balance the load on different data centers and VM (Virtual Machine) according to the tasks. The result shows that the proposed algorithm performs better in resource allocation and minimize overall propagation time of input and output data. The effect of increasing the load on the server with and without applying the proposed load balancing algorithm is analysed. The proposed load balancing algorithm is designed to decreases the response time by the server. The user request is processed virtually in different cloud, and achieves the final goal.

The round robin algorithm is used for load distribution. This method has proven quite effective in real world scenarios. The round robin algorithm can be effective for distributing the workload among servers with equal processing ability. When servers differ in their processing capacity, by response times or amount of active relations as the selection criteria can optimize user response times. This work is basically an extension of Round Robin (RR), Throttled Load Balancing Algorithm (TLBA), and Equally Spread Current Execution Algorithm (ESCEA). With round robin scheduling algorithm, there is a limitation of same maps or mob for each and every cloudlet and it takes each task in round robin order. The CloudSim simulator is used to simulate cloud environment for experimentation.

A. WORK FLOW

Service Proximity Policy: The broker selects the data center according to round robin scheduling and allocates the cloudlets with different maps and mobs to the selected data center. Allocation of cloudlets required following steps.

Begin Main

Step1: Start the CloudSim Simulator

Step2: Initialize data centers with their configurations i.e. their processing speeds, RAM, bandwidth etc.

Step3: Register each data center with CIS and in turn that will report to the main broker for handling cloudlets.

Step4: Then broker will receive the cloudlet requests.

Step5: Based on the data centers chosen in round robin order broker allocated the received cloudlets to different ate centers.
Step6: Repeat step 4 and 5 until no more cloudlets.
Step7: End Main

Table 7: Representation of load balancing algorithm on Cloud analyst for Cost Estimation and Average Requesting Time Table

Parameters	***Load Balancing Algorithm on Cloud Analyst*** ***Round Robin*** ***ESCE*** ***Throttled***		
Data Center's	*2*	*4*	*3*
User Base	*5*	*2*	*4*
H/W Unit	*2*	*7*	*6*
Virtual Machine	*20*	*25*	*30*
Average(ms)	*0.28*	*0.48*	*0.38*
Minimum(ms)	*0.02*	*0.04*	*0.06*
Maximum(ms)	*0.64*	*0.08*	*0.54*
Total $ (Cost)	*57.66*	*38.60*	*43.98*

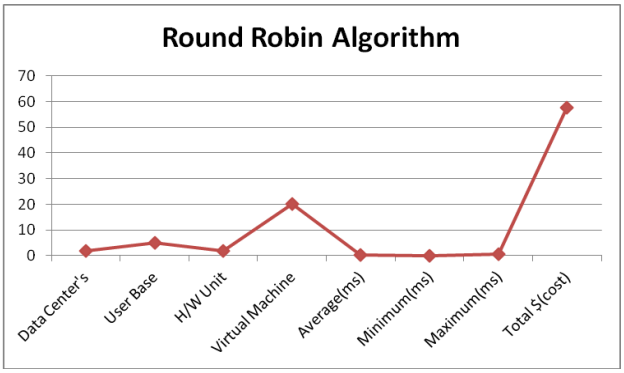

Fig 5: Performance of Round Robin Algorithm

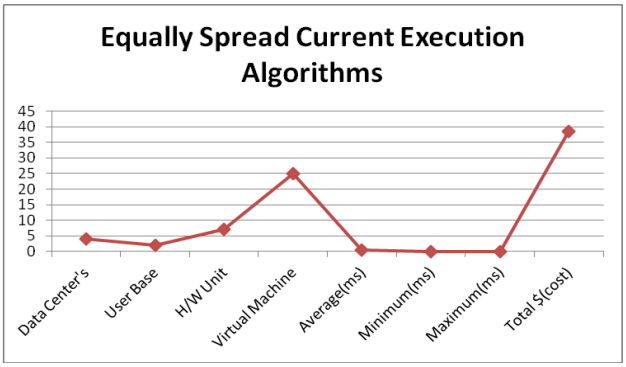

Fig 6: Performance of equally spread current Execution Algorithm

The performances of Round Robin Algorithm and Equally Spread Current Execution Algorithm on the load balancing in cloud are illustrated in Fig 5 and 6.

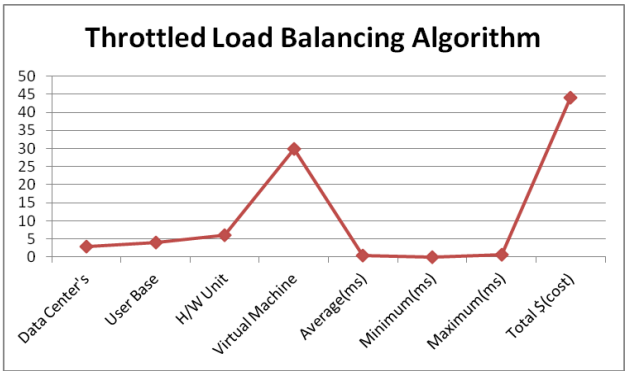

Fig 7: Throttled Load Balancing Algorithm

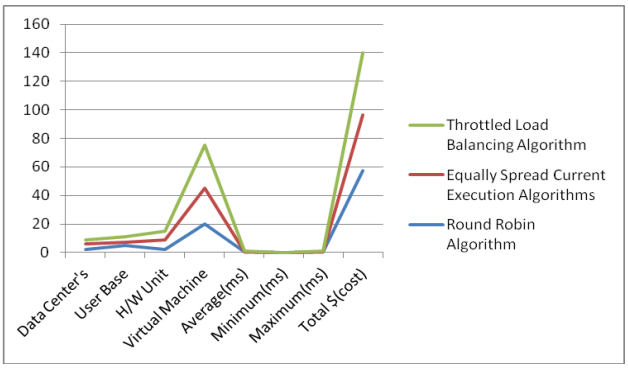

Fig 8: Comparison of Three Algorithms

Fig 7 shows the graphical representation of the data centers, user base, h/w unit, virtual machine, average (ms), minimum (ms), maximum (ms) and total of the Overall Response Time for TLBA. The comparisons of the overall RBA, ESCE and TLBA are represented in Fig 8.

VII. CONCLUSION

In this paper, the response time and the dynamic resource allocation are discussed, which are the growing need of cloud providers to provide services for more number of users. Dynamically scalable virtualized resources are discussed. Recent computers are sufficiently powerful to use virtualization to present the deception of many smaller virtual machines. This work surveyed, various load balancing techniques for cloud computing. In the surveyed literatures, different dynamic intelligent load balancer were proposed and implemented in Cloud. Cloud services are simpler to acquire and scale up and down. The future work will be on the development of more effective allocation algorithm in heterogeneous nature and works in dynamic environment using virtual Machines.

REFERENCES

[1] M. Mezmaza, N. Melba, Y. Kessaci b, Y.C. Lee, E. G. Talbi, A.Y. Zomayac, D. Tuyttens, A parallel biobjective hybrid met heuristic for energy-aware scheduling for cloud computing systems, Journal of Parallel and Distributing Computing, 2011.

[2] J. T. Hung, and T. G. Robertazzi. Scheduling non linear computational loads. IEEE Transactions on Aerospace and Electronic Systems, 2008.

[3] T. Sridhar, Cloud Computing: A Primer, Part 1: Models and Technologies. The Internet Protocol Journal, 2009.

[4] Soto mayor, B., RS. Montero, M. Llorente, and I. Foster, "Virtual infrastructure management in private and hybrid clouds," in IEEE Internet Computing, 2009.

[5] Radojevic, B. and M. hZagar, "Analysis of issues with load balancing algorithms in hosted (cloud) environments." In proc 34th International Convention on MIPRO, IEEE, 2011.

[6] Wang, S-C., K-Q. Yan, W-P. Liao and S-Swung, "Towards a load balancing in a three-level cloud computing network," in proc. 3rd International Conference on. Computer Science and Information Technology (ICCSIT), IEEE, July 2010.

[7] Hamada, M., & Tahvildari, L. (2012). Cloud Computing Uncovered: A Research Landscape. In H.Ali & M. At if (Eds.), Advances in Computers, 2012.

[8] Shiraz M, Agni A.A Light weight Active Service Migration Framework for Computational Offloading in Mobile Cloud Computing Journal of Supercomputing 2013.

[9] Indrajit Roy, Srinath T.V. Settee, Ann Kilzer, Vitally Shmatikov, Emmett Witches, "Air vat: Security and privacy for map reduce", in the proceedings of the 7th USENIX conference on networked systems design and implementation, 2010.

[10] T-Y., W-T.Lee, Y-S.Lin, Y-S.Lin, H-L.Chan and J-S. Huang, "Dynamic load balancing mechanism based on cloud storage" in proc. Computing, Communications and Applications Conference (Com Com Ape), IEEE, pp: 102-106, January 2012.

[11] Ellinger, R. S. (2013). Governance in SOA Patterns (white paper).In The Northrop Grumman Corporation for Consideration by OASIS SOA Reference Architecture Team, 2013.

[12] D. Escalate, Andrew J. Borty, "Cloud Services: Policy and Assessment", Educes review July/August 2011.

[13] Tushar Desai, Jignesh Prajapati "A Survey of Various Load Balancing Techniques and Challenges in Cloud Computing" International Journal of Scientific & Technology Research Volume 2, Issue 11, November 2013.

[14] M.Aruna, D.Bhanu, R.Punithagowri "A Survey on Load Balancing Algorithms in Cloud Environment" International Journal of Computer Applications, Volume 82 – No 16, November 2013.

[15] R. X. T. and X. F. Z., A "Load Balancing Strategy Based on the Combination of Static and Dynamic, in Database Technology and Applications (DBTA)", 2010 2nd International Workshop, pp. 1-4, 2010.

[16] Liang-Teh Lee, Kang-Yuan Liu, Hui -Yang Huang and Chia- Ying Tseng, "Dynamic Resource Management with Energy Saving Mechanism for Supporting Cloud Computing" International Journal of Grid and Distributed Computing, Feb 2013.

[17] Tushar Desai, Jignesh Prajapati, "A Survey of Various Load Balancing Techniques and Challenges in Cloud Computing" International Journal of Scientific & Technology Research, 2012

[18] Soumya Ray and Ajanta De Sarkar "Execution Analysis of Load Balancing Algorithms in Cloud Computing Environment" International Journal on Cloud Computing: Services and Architecture (IJCCSA), Vol.2, No.5, October 2012.

[19] L. Doddini Probhuling "Load Balancing Algorithms in Cloud Computing" International Journal of Advanced Computer and Mathematical Sciences ISSN 2230-9624 and 2013.

[20] Shivam Tundele, Anshul Maru "Implementation of Network Load Balancing System" International Journal of Engineering Sciences & Research Technology ISSN: 2277-9655,and February 2014.

[21] Pooja Samal, Pranati Mishra "Analysis of variants in Round Robin Algorithms for load balancing in Cloud Computing" International Journal of Computer Science and Information Technologies, ISSN: 0975- 0982, and 2013.

[22] Nusrat Pasha, Dr. Amit Agarwal "Round Robin Approach for VM Load Balancing Algorithm in Cloud Computing Environment " International Journal of Advanced Research in Computer Science and Software Engineering, ISSN: 2277 1288K, and May 2014.

[23] M. Durairaj, A. Menaka, "A Survey of Soft Computing Approach for Load Balancing in Cloud Computing" International Journal of Emerging Technology and Innovative Engineering, ISSN: 2394 – 6598 and February 2015.

