International Journal of Trend in Research and Development, Volume 5(3), ISSN: 2394-9333
www.ijtrd.com

Taxonomy of Load Balancing Techniques in Cloud Computing

[1]P.Ratna Shekar and [2]A.Sandhya,
[1,2]Assistant Professor, Department of Computer Science, University P.G.College, Secunderabad, Telangana State, India

***Abstract*—** Day by day progression of IT industry is advancing swiftly and to meet its advanced requirements there is need of highly reliable, fast, and scalable computing and storage services. Cloud computing is a computing technology that helps organizations to make use of computing and storage services over internet in paid basis. From past decade popularity of cloud computing is increasing and most of the organizations are looking forward to migrate to cloud computing because of its diversified applications such as reliability, scalability, agility, and moreover it reduces the burden of cost and IT infrastructure administration related issues. Cloud computing is a distributed computing environment and is in demand for its service delivery and efficient outputs. But there are also several issues like security, load balancing, etc. exists, that are need to be resolved for better outcome of the business. Load balancing plays significant role in enhancing performance of cloud computing environment. Distributing the load between several resources for better usage and increasing the response time by avoiding jobs that consumes excess usage of resources while some other tasks are idle and waiting for those resources. Here load means tasks of particular resource like CPU, storage, network etc. to be executed. There are several techniques are proposed for the handle the load balance issue in cloud computing. This paper discusses some of those techniques which helps cloud to improve the reliability and produce better availability of resources.

***Keywords*—** *Cloud Computing, Load Balancing, Load Balancing Techniques, Load Balancing Advantages, Cloud Resources.*

I. INTRODUCTION

Cloud computing provides virtualized resources and administrations with prospective of reducing expenses. It became prominently popular technology nowadays for the reason that it has characteristically numerous advantages [21]. In cloud computing clients access resources of cloud server through paid registration and by sending requests to access the resources. After verifying the client's credentials cloud service providers permits clients to access virtualized the resources. These resources can be storage, operating systems, network, and software applications etc [25]. Cloud computing offers its services in three service layer and four deployment models. The three service layers are Infrastructure as a Service layer (IaaS) in which hardware infrastructures like storage, network etc. are provided. SaaS (Software as a Service) layer provides software applications and PaaS (Platform as a Service) provides computing platforms to clients for developing their own software applications. Public, Private, Community, and Hybrid models are cloud deployment models. In public cloud services are accessible to general public, in private models services are accessible only to particular organization, in community model services are offered to specific communities and in hybrid cloud model is combination of one or cloud models. The major challenge in cloud computing is to meet the clients demands for acceding cloud resources. Cloud servers needs to handle the traffic and workload requested by client. Here load balancing plays crucial role. Load balancing cloud computing is method of allocating workloads and resources over more than servers in such a way that it promises the maximization of productivity with lesser response time. Primary objective of cloud load balancing is to maintain business continuity, maintaining of system efficiency, enhancing the performance, maintain the availability of the resources against system failures. The load balancer dispenses client requests to various servers according to which server or node is currently engaging other requests. Without load balancing, clients must wait until the previous requests are processed by server, this situation increases response time and traffic or load on server. Before allocating server load balancing process checks request waiting in queue, CPU processing rate, time of request placed etc. Figure 1 shows typical cloud load balancing process. Load balancing has various advantages [32]. With Cloud load balancing techniques are easy to put into practice and cost effective. With load balancing organizations produce high performing applications by making their applications work faster and offer enhanced performance with lesser costs. With help of cloud characteristics like agility and scalability cloud load balancing handles traffic. Efficient load balancing techniques effectively handles the huge traffic and client distributes requests over several servers.

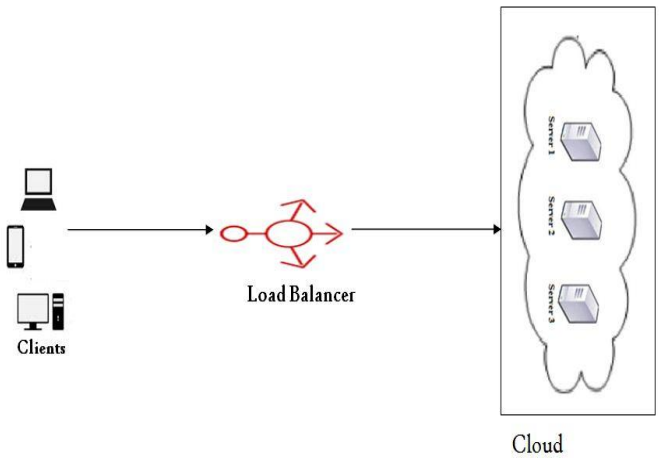

Fig. 1. Cloud Computing Load Balancing Process

With cloud load balancing techniques servers can easily handles excessive client requests at the same time and irrespective of size of the request, it can be efficiently distributes over different servers in order to get significantly improved results in lesser response times. The basic objective of load balancing techniques is to save the rescue system in unforeseen outages. With load balancing distributing workload over several servers although one server fails, another active server will be ready to handle the workload. Generally load balancing techniques are categorized into four categories based on current system status: one is static and dynamic methods.

Static load balancing techniques are non pre-emptive and distributes requests based on previous client requests data like computing, network etc. Dynamic load balancing techniques redistributes resource requests between several processors at the time of execution. Based on decision approach load balance techniques are categorized into centralized, distributed and hierarchical load balancing algorithms. In centralized technique single node server is acts central node and allocates and schedules resources. This central node has all the details of entire cloud network and performs static or dynamic load balancing. This technique offers less response time but does not provides fault tolerance. In distributed load balancing, allocation of resources if processed by multiple servers rather than a single node, each node contains information that is used to distribute tasks with static and/or dynamic load balancing. Hierarchical load balancing adopts tree model layered structure in such a way the each node in the tree is maintained by parent node. Parent node collects the data from child node by applying lighter processing. Next section discusses various load balancing techniques.

I. LOAD BALANCING TECHNIQUES

A. Round Robin

Round robin is fundamental load balancing techniques allocates the resources to latest requests that are circulated among available servers. The main advantage of this technique is that it is simple and easy to implement. Disadvantage of this technique is that it uses previous tasks information to allocate resources and doesn't use the present system information.

B. Throttled Load Balancing

Throttled Load Balancing is a dynamic load balancing technique, where clients place their requests to a control center and maintains current state of system. This control center requests virtual machine load balancer for deciding the suitable virtual machine to take control of the workload. Throttled load balancer maintains a list of virtual machines and their idle or busy status. Client requests are allocated when appropriate virtual machine available to handle that request which is made by control center. If the suitable virtual machine is not available, clients need to wait until it becomes available.

C. RASA

RASA (Resource Aware Scheduling Algorithm) is load balancing technique that allocates the server to requests by using algorithms Min-Min and Max-Min. Requests resource execution time is estimates by these Min-Min and Max–Min algorithms to use the servers based on execution time and this technique gives enhanced results by doing this.

D. ESCE VM load Balancing

ESCE or Equally Spread Current Execution technique is refereed as Active Virtual Machine Load Balancing technique which uses spread spectrum method and uniformly allocates workload between servers of cloud data center. Request remains in queue until all the client requests that are executed by VM. This technique frequently verifies the requests that are in queue and list virtual machines that are processing other requests. Request is allocate the virtual machine that is free and if any virtual machine handling excessive requests, transfers the workload to other idle virtual machine. Disadvantage of this technique is that it requires heavy computational workload.

E. AMLB Technique

This technique keeps track of each VM and different requests currently handled by the VM in the form of a table. When request comes, AMLB checks for the virtual machines which are handling lesser requests. If it found the virtual machine it sends virtual machines id to control center in order to handle request by that virtual machine

F. Honey Behavior Technique

In honey behavior load balancing technique request that have higher priority are released from heavy loaded virtual machines and allocated to virtual machines that have least burden of requests. Profitable virtual machine that has least process execution time is also balanced with assignment of reliable tasks in a pre-emptive way. This technique can achieve less execution time and makes better use of resources.

G. Active Clustering

Active Clustering is a load balancing technique that groups alike nodes and operates based on those grouped nodes. Initially a node make initiation for handling requests and picks other node to transfer request from node (go-between node) to its neighbor node which satisfies the factors that are unique than the priori nodes. This go-between node makes link among one of the neighbor node that has similar characteristics. After that go-between node eliminates the link between similar node and itself. This process is carried out until the load balancing is achieved.

H. PA-LBIMM Technique

The PA-LBIMM (Priority Aware Load Balancing Improved Min Min Algorithm) splits the requests into G1 and G2 groups. Such that requests submitted by clients who have higher priority are placed into group G1 and requests submitted by low priority clients are placed into group G2. Request processed into to resources according to their priority. Fore-mostly all the requests that are placed under G1 are processed using Min-Min algorithm and after the every request in group G2 processed by using Min-Min algorithm.

II. CONCLUSION

Load balancing has crucial impact on the performance of cloud computing. The main aim of load balancing techniques is to make productive resource usage, business continuity, availability of resources, maximum throughput, minimizing response time, increasing reliability and stay away from overloading of single resource. Good load balancing technique helps cloud computing to become more productive and improves client's gratification. This paper surveys some of those popular load balancing techniques that are static, dynamic, composite, and prioritized.

References

[1] D. Suresh Kumar, Dr. E. George Dharma Prakash Raj, "A Literature Review On Load Balancing Mechanisms In Cloud", Ijarcs, Volume 9, No. 1, January-February 2018.

[2] Davneet Singh Chawla, Dr. Kanwalvir Singh Dhindsa, "A Load Balancing Based Improved Task Scheduling Algorithm In Cloud Computing", Ijraset, Volume 5, Issue 9, September 2017.

[3] S. Mubarak T., V Rangesh, Dr G.Umarani Srikanth, "A Survey On Load Balancing In Cloud Computing Using Optimization Technique". Ijre, Volume 04, Issue 01, January 2017.

[4] Shashi Mehar, Hansa Acharya, "A Survey On Modern Load Balancing Algorithms In Cloud Computing", Ijer, Volume 6, Issue No.4, April 2017.

[5] G. Kanmani, E Jayabalan, "A Survey On Recent Improved Load Balancing Algorithms In Cloud Environment", Ssrg International Journal Of Computer Science And Engineering – Ncsact – 2017.

[6] Avnish Thakur, Major Singh Goraya, "A Taxonomic Survey On Load Balancing In Cloud", Jnca, August 2017.
[7] Rajpreet Kaur And Vikas Khullar, "Abc Optimized Weighted Task Load Balancing Algorithm In Cloud Computing", Iijcs, Volume 5, Issue 9, September 2017.
[8] Er. Imtiyaz Ahmad , Er. Shakeel Ahmad, Er. Sourav Mirdha, "An Enhanced Throttled Load Balancing Approach For Cloud Environment", Irjet, Volume 04, Issue 06, June 2017.
[9] Meenakshi Sharma, Anitha Y And Pankaj Sharma, "An Optimistic Approach For Load Balancing In Cloud Computing", Ijca, Volume 2, Issue 3, March 2014.
[10] Jyoti Rathore, Dr. Bright Keswani, Dr. Vijay Singh Rathore, "Analysis Of Various Load Balancing Techniques In Cloud Computing: A Review", Suresh Gyan Vihar University Journal Of Engineering & Technology, Vol . 3, Issue 2, 2017.
[11] Zenon Chaczko, Venkatesh Mahadevan, Shahrzad Aslanzadeh And Christopher Mcdermid, "Availability And Load Balancing In Cloud Computing", Ipcsit, Volume 14, 2011.
[12] Suman Pandey, "Cloud Load Balancing: A Perspective Study", Ijecs, Volume 6, Issue 6, June 2017.
[13] Trapti Gupta, Abhishek Dwivedi, "Data Storage & Load Balancing In Cloud Computing Using Container Clustering", Ijess, September, 2017.
[14] Samriddhi Gupta, Abhishek Bajpai, Dr. Neeraj Kumar Tiwari, "Dynamic Load Balancing In Cloud Computing Using Ant Colony Optimization", Ijsrset, Volume 3, Issue 3, 2017.
[15] Shreenath Acharya, Dr. Demian Antony D'mello, "Energy And Cost Efficient Dynamic Load Balancing Mechanism For Resource Provisioning In Cloud Computing", International Journal Of Applied Engineering Research, Volume 12, Number 24, 2017.
[16] M. Nirmala, Dr. T.Adi Lakshmi, "Energy-Efficient Load Balancing In Cloud: A Survey On Green Cloud", Ijett, Special Issue, April 2017.
[17] S. Subalakshmi, N. Malarvizhi, "Enhanced Hybrid Approach For Load Balancing Algorithms In Cloud Computing", Ijsrcseit, Volume 2, Issue 2, 2017.
[18] Mahfooz Alam And Zaki Ahmad Khan, "Issues And Challenges Of Load Balancing Algorithm In Cloud Computing Environment",Indian Journal Of Science And Technology, Volume 10, July 2017.
[19] T. Deepa, Dr. Dhanaraj Cheelu, "Load Balancing Algorithms In Cloud Computing: A Comparative Study", Ijiacs, Volume 6, Issue 1, January 2017.
[20] Asma Anjum, Dr. Rekha Patil, "Load Balancing For Cloud Ecosystem Using Energy Aware Application Scaling Methodologies", Irjet, Volume 04, Issue 05, May 2017.
[21] Ruchika Aggarwal, Latika Gupta, "Load Balancing In Cloud Computing", Ijcsmc, Volume 6, Issue 6, June 2017.
[22] Navpreet Singh, Dr. Kanwalvir Singh Dhindsa, "Load Balancing In Cloud Computing Environment: A Comparative Study Of Service Models And Scheduling Algorithms", Int. J. Advanced Networking And Applications, Volume 08, Issue 06, 2017.
[23] Nikhil Rajeshirke, Rohan Sawant, Sumeet Sawant, Hasib Shaikh, "Load Balancing In Cloud Computing", Ijrter, 2017.
[24] Sambit Kumar Mishra, Bibhudatta Sahoo, Priti Paramita Parida, "Load Balancing In Cloud Computing: A Big Picture", Journal Of King Saud University - Computer And Information Sciences, January 2018.
[25] Acharya Mitali Nilesh, Prof. Chirag A. Patel, "Load Balancing In Cloud Computing Using Ant Colony Optimization", Ijcet, Volume 8, Issue 6, Nov-Dec 2017.
[26] Shweta Patel, Prof. Mayank Bhatt, "Load Balancing In Cloud Computing Using Cloudsim", Ijrise, Volume 3, Issue 2, March-April 2017.
[27] Rajwinder Kaur And Pawan Luthra, "Load Balancing In Cloud Computing", Association Of Computer Electronics And Electrical Engineers, 2014.
[28] Suman Rani, Vinod Saroha, Sanjeev Rana, "A Hybrid Approach Of Round Robin, Throttle & Equally Spaced Technique For Load Balancing In Cloud Environment", Ijiacs, Volume 6, Issue 8, August 2017.
[29] Alireza Sadeghi Milani, Nima Jafari Navimipour, "Load Balancing Mechanisms And Techniques In The Cloud Environments: Systematic Literature Review And Future Trends", Journal Of Network And Computer Applications, June 2016.
[30] Einollah Jafarnejad Ghomi, Amir Masoud Rahmani, Nooruldeen Nasih Qader, "Load-Balancing Algorithms In Cloud Computing: A Survey", Journal Of Network And Computer Applications, April 2017.
[31] Harish Chandra, Himanshu Bahuguna, "A Survey Of Load Balancing Algorithms In Cloud Computing", Ijcea, Volume Xi, Issue Xii, Dec. 17.
[32] Klaithem Al Nuaimi, Nader Mohamed, Mariam Al Nuaimi And Jameela Al-Jaroodi, "A Survey Of Load Balancing In Cloud Computing: Challenges And Algorithms", Ieee Second Symposium On Network Cloud Computing And Applications, 2012.
[33] Priya Abraham, "What Is Load Balancing In Cloud Computing And What Are Its Advantages?", Article In Znetlive.Com Blog, 29 May 2017.