Volume 4, Issue 2, February 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Review on Existing Load Balancing Techniques of Cloud Computing

Ruhi Gupta

Department of Computer Science, PURCITM

Punjabi University, Patiala, India

Abstract: Cloud Computing is an emerging paradigm of the computing world which aims to share data over a single platform. Since 2007, cloud computing has emerged as a buzzwords in ICT industry. Two players in cloud computing environments, cloud providers and cloud users, pursue different goals; providers want to maximize revenue by achieving high resource utilization, while users want to minimize expenses while meeting their performance requirements. However, it is difficult to allocate resources in a mutually optimal way due to the lack of information sharing between them. As an increased demand of the resources of cloud computing, load balancing is the usual problem to be faced. Various load balancing algorithms have been designed by various researchers. A close look at those algorithms and the comparison between those algorithms is being done in this review paper.

Keywords: Cloud Computing, Load Balancing, Honey Bee Foraging Algorithm, Ant Colony Optimization Technique, RBAC, RASA

I. Introduction:

Cloud Computing is a new era which aims to have shared data over a one platform. As the technology is booming fast, so does the requirements of the clients. This new paradigm of cloud computing is appealing vendors which increases its popularity. As the definition of NIST[9] says "Cloud Computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g. network, server, storage and applications and services) that can be rapidly provisioned and released with minimal management efforts or service provider interaction.

Cloud Computing Services:

According to NIST, Cloud Computing provides 3 service models that are described in the table(1) below and these three models club to form cloud computing architecture.

Table 1: Services offered by Cloud Computing

Services offered	Definition	Layer	E.g.
1. SaaS (Software as a Service)	SaaS provides the vendor with the software. Vendor pays for the time of using the software and can use it anywhere. There is no need to buy the software. It is referred to as "on-demand" software.	Application Layer	GoogleApps, salesForce.com
2. PaaS (Platform as a Service)	It provides a platform where resources are available and consumers can themselves create the required applications for e.g. web application data and database data. Providers provide network, servers, and storage services.	Platform Layer	Google AppEngine, Microsoft Azure
3. IaaS (Infrastructure as a Service or Hardware as a Service)	This model provides users with the hardware on rent for e.g. server space, network equipment, memory, storage space. In short consumer buys virtual space and works on it.	Hardware Layer	Amazon EC2, Simple Storage Service(S3)

II. Load Balancing:

Load Balancing is a technique to distribute the load evenly among all the nodes of the network. If any node is heavy i.e. have more load than required then its load is given to the node with less load. Hence load balancing helps the overloaded and under loaded nodes. Load balancing[5] is a major challenge of cloud computing. Goals of load balancing involve:

1. Optimum resource utilization
2. Maximum throughput
3. Maximum response time
4. Avoiding overload

Load balancing can be static or dynamic.

a) Static load balancing: In this, load is distributed evenly across all the nodes. This algorithm must know about the system resources beforehand.

b) Dynamic load balancing: It depends on the current state of the system. If any node is overloaded then its load is shifted to the under loaded node. So real time communication is performed here.

Fig.1 shows the structure of cloud computing after load balancing algorithms implementation.

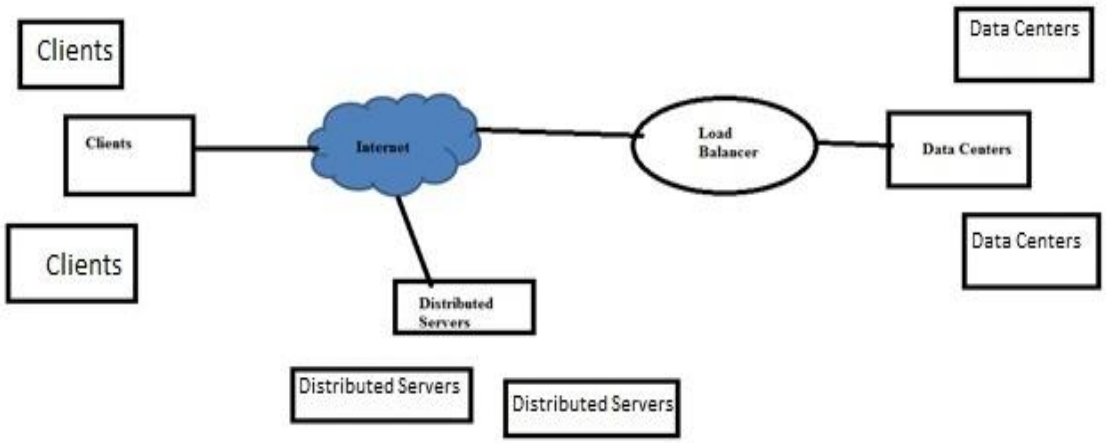

Fig.1 : Load Balancing in Cloud Computing

A. Policies followed in dynamic load balancing are:

Following are the policies of load balancing[5]:

-Transfer Policy: The part of the dynamic load balancing algorithm which selects a job for transferring from a local node to a remote node is referred to as Transfer policy or Transfer strategy.

-Selection Policy: It specifies the processors involved in the load exchange (processor matching)

- Location Policy: The part of the load balancing algorithm which selects a destination node for a transferred task is referred to as location policy or Location strategy.

-Information Policy: The part of the dynamic load balancing algorithm responsible for collecting information about the nodes in the system is referred to as Information policy or Information strategy.

Fig.2 shows how all these policies works.

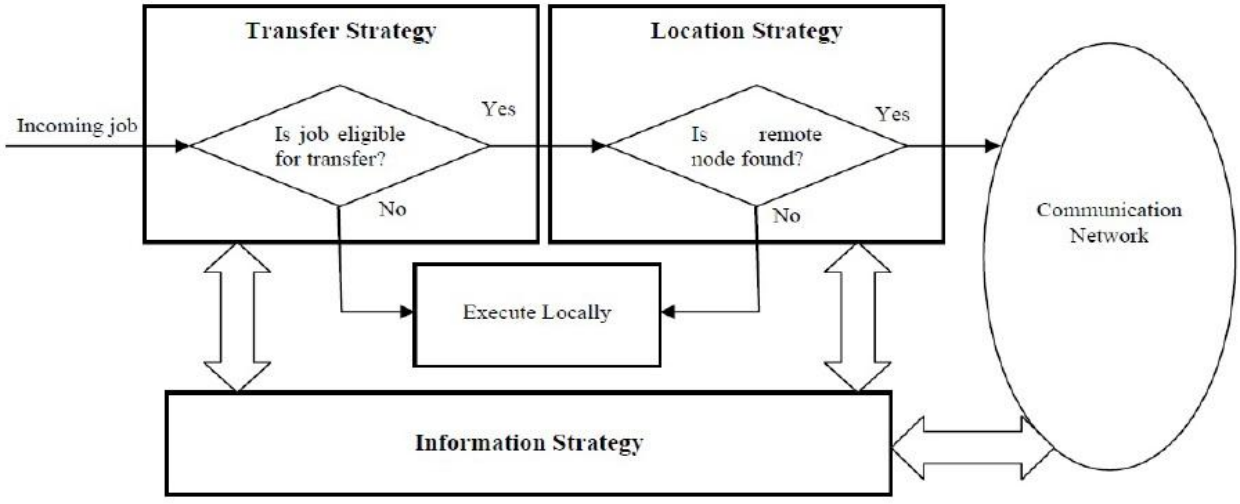

Fig.2 Policies Implementation[5]

III. Existing Load Balancing Techniques:

There are various techniques[1],[3],[4],[6] to balance the load of cloud computing. Some of which are discussed in this paper.

1. **Honey Bee Foraging Algorithm**: This whole algorithm is based on the process of honeybees finding the food and alarming others to go and eat the food. First forager bees go and find their food. After coming back to their respective beehive, they dance. After seeing the strength of their dance, the scout bees follow the forager bees and get the food. The more energetic the dance is, the more food available is. So this whole process is mapped to overloaded or under loaded virtual servers. The server processes the requests of the clients which is similar to the food of the bees. As the server gets heavy or is overloaded, the bees search for another location i.e. client is moved to any other virtual server. In this way, this whole technique works.
2. **Task Scheduling Algorithm based on Load Balancing:** Y. Fang et al.[10] discussed a two-level task scheduling mechanism based on load balancing to meet dynamic requirements of users and obtain high resource utilization. It achieves load balancing by first mapping tasks to virtual machines and then virtual machines to host resources.
3. **Throttled Load Balancing Algorithm**: This algorithm makes use of identity of virtual machines. Client requests the ID of virtual machine. Throttled load balancing algorithm returns that ID to the user.
4. **Ant Colony Optimization Technique**: In this technique, a pheromone table was being designed which was updated by ants as per the resource utilization and node selection formulae. Ants move in forward direction in search of the overloaded or under loaded node. As the overloaded node is traversed, then ants move back to fill the recently encountered under loaded node, so a single table is updated every time.
5. **Role Based Access Control (RBAC)**: RBAC is a technique used to reduce the load of the cloud. In this, a role is assigned to each user so that limited applications of the cloud can be accessed by their respective number of users. So by this approach, the resources are restricted to the users.
6. **Resource Allocation Scheduling Algorithm (RASA):**In this algorithm, virtual nodes are created first. Then the expected response time of each virtual node is found. Then according to the least loaded node criteria, efficient virtual node is found and ID of that node is returned to the client. In this, Min-Min and Max-Min strategies are followed. If number of resources available are odd, then Min-Min strategy is applied else Max-Min strategy is applied.

   Comparison between all these techniques is shown in Table 2.

Table 2:Comparison of the above discussed techniques

Load Balancing Methods	*Parameters*	*Merits*	*Demerits*
1. Honey Bee Foraging Technique	Execution Time, Overheads, Throughput	Maximize throughput, Low overheads	Low Priority Load
2. Task Scheduling Algorithm based on Load Balancing	Response Utilization, Performance, Response to Request Ratio	Maximise Response Utilization, Increased Performance	Doesn't improve response to request ratio
3. Throttled Load Balancing Algorithm	Load Movement Factor, Communication Cost, Network Delay	High Load Movement Factor	High Communication Cost, High Network Delay
4. Ant Colony Optimization	Performance, Resource Utilization, Fault tolerance, Scalability	Performance increased, Resource utilization high, Fault tolerance excellent, Scalability Good	Complex Network
5. RBAC	Performance, Resource Utilization, Energy, CPU Burst Time	High Performance, High Resource Utilization, CPU Burst time decreases	Response Time increases
6. RASA	Performance, Execution Time	Performance increases, Execution time decreases	Less Fault Tolerance

IV. Conclusion:

As the Cloud Computing is an alluring concept in the present and upcoming time, the researchers had developed various techniques to cope up with the load balancing problem being faced while working with cloud computing. In this paper a light is being thrown on the recent techniques being used. And comparison is done between them based on various parameters. Further work can be done to balance the load by distributed policy so as to reduce the workload of the executer.

References:

[1] R.G.Rajan , V.Jeyakrishnan, "A Survey on Load Balancing in Cloud Computing Environments", *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 2, Issue 12, December 2013

[2] P. Salot, "Asurvey of various scheduling algorithm in cloud computing environment*", IJRET*, Volume: 2 Issue: 2, Feb 2013

[3] A.K. Sidhu, S. Kinger, "Analysis of Load Balancing Techniques in Cloud Computing", *International Journal of Computers & Technology*,Volume 4 No. 2, ISSN 2277-3061, March-April, 2013

[4] S. Mohana Priya, B. Subramani," A new approach for load balancing in cloud computing*", International Journal of Engineering and Computer Science*, ISSN: 2319-7242 Volume 2 Issue 5 May, 2013 Page No. 1636-1640

[5] Suresh M., Shafi Ullah Z., Santhosh Kumar B.," An Analysis of Load Balancing in Cloud Computing*", International Journal of Engineering Research & Technology (IJERT)*,Vol. 2 Issue 10, October – 2013, ISSN: 2278-0181

[6] A. Kaur,N. Bansal, "Cross-Breed Job Scheduling for Reducing the Server Load Using RBAC at Cloud*", International Journal of Advanced Research in Computer Science and Software Engineering*", Volume 3, Issue 5, ISSN: 2277 128X, May 2013

[7] N. J. Kansal, I. Chana," Cloud Load Balancing Techniques: A Step towards Green Computing*", IJCSI International Journal of Computer Science Issues*, Vol. 9, Issue 1, No 1, January 2012, ISSN (Online): 1694-0814

[8] N. Sran, N. Kaur, "Comparative Analysis of Existing Load Balancing Techniques in Cloud Computing*", International Journal of Engineering Science Invention* ,Volume 2 Issue 1, PP.60-63,ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726, January. 2013

[9] P. Mell,T.Grance, "The NIST Definition of Cloud Computing (Draft)"

[10] Y. Fang, F. Wang, and J. Ge, "A Task Scheduling Algorithm Based on Load Balancing in Cloud Computing", *Web Information Systems and Mining*, *Lecture Notes in Computer Science*, Vol. 6318, pages 271-277,2010.

