ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 9, September 2014

A Survey on Load Balancing in Cloud Computing Using Soft Computing Technique's

Alok singh[1], Vikas Kumar Tiwari[2], Dr. Bhupesh Gour[3]

PG Scholar,DoCSE T.I.T, Bhopal, India[1]

Assistant Professor, DoCSE,T.I.T, Bhopal, India[2]

Additional Director, DoCSE,T.I.T, Bhopal, India[3]

Abstract: As the IT industry growing in day to day, the need of computing and storage is increasing sharply. Exchange of data over the network is also continuously increasing. New advance technology, cloud computing is become popular because of providing above services beneficially. Other vital technologies such as virtualization and scalability by creating virtual machines in cloud computing. In cloud computing web traffic and service provisioning is increasing day by day so load balancing is becomes a big research issue in cloud computing. Cloud Computing is a new propensity emerging in IT environment with in huge requirements of infrastructure and resources. Load Balancing technique for cloud computing is a vital aspect of cloud computing environment. Adept load balancing scheme ensures adept resource utilization by provisioning resources to cloud users on-demand services basis in pay-as-you-use manner. Load Balancing may further support prioritizing users by applying appropriate scheduling criteria. In this survey paper presents various load balancing schemes in different cloud computing environment basis on requirements specified in Service Level Agreement (SLA) process and Load balancers are used for assigning load to various virtual machines in such a way that none of the nodes gets loaded heavily or lightly. The load balancing needs to done properly because failure in any one of the node can lead to unavailability of data.

Keywords: cloud computing , load balancing , SLA , virtual machine , load balancing algorithm.

INTRODUCTION

Due to the biggest success of internet in last few years, computing resources is now more everywhere available And it enabled achievement of a new computing concept called Cloud Computing. Cloud Computing environment necessitate the traditional service providers to have two various ways. These are infrastructure and service providers. Infrastructure providers arrangement of cloud platforms and lease resources according to usage. Service providers provide a resources from infrastructure providers to serve end users. Cloud Computing has tempt the giant companies, like Google, Amazon and Microsoft considered as a great effect in today's Information Technology companies. cloud computing concept attracted in several feature .These are as follows:-

- Lower initial investment
- Easier to manage
- Scalability
- Deploy faster
- Location independent
- Device independent
- Reliability
- Security

despite the fact that cloud computing has shown Enough opportunities to the IT industry of today's world, but still there are number of challenges that requires to be carefully addressed. Our aim is provide a best understanding of cloud computing and focus on the research ongoing in this hugely flourishing amphitheatre of computer science.

2. CLOUD COMPUTING OVERVIEW

2.1 What Is Cloud Computing?

Cloud computing use a technology for the internet and central remote servers to maintain applications and data. Cloud computing allows consumers and businesses to use applications without installation and access their personal files at any computer with internet access. This technology allows for much more inefficient computing by centralizing storage, memory, processing and bandwidth. Cloud computing is a model of network computing where a program or application runs on a connected server or servers rather than on a local computing device such as a PC, tablet or smart phone. Like the Conventional client-server model or chronic mainframe computing, a user connects with a server to execution a task. The difference with cloud computing is that the computing process may run on one or many connected computers can be, utilizing concept of virtualization. With virtualization is one or more physical servers can be configured and splits into multiple unattached "virtual servers, all functioning independently and seem to the user to be a single physical device. Such virtual servers is do not physically consist and can therefore be moved in all directions and flaky up or down on the fly without affecting the end user. The computing resources have become " grainy ", which endow end user and operator avail including broad access

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 9, September 2014

across multiple devices, resource pooling, on-demand service, Fast elasticity and service surveying capability.

Cloud computing is a mechanism of distributed computing that focuses on confer a wide range of users with distributed access to virtualized hardware and software infrastructure over the internet. It involves distributed computing virtualization, networking, web services and we software. idea of cloud computing has focus interest of users towards of parallel , distributed and virtualization computing systems today. It has appear as a popular solution to provide cheap and easy access to externalized IT resources .Through virtualization, cloud computing is able to address with the same physical infrastructure a large client base with different computational needs. The rapid growth in the field of cloud computing also increases severe security concerns. Lack of security is the only hurdle in wide adoption of cloud computing.

2.2 Types Of Cloud

2.2.1 Public cloud In Public cloud is available for public use alternatively for a large industry and is owned by an organization selling cloud services . customer has no visibility and control Excess where the computing infrastructure is hosted. The computing infrastructure is shared among any organizations**.**

22.2.Private cloud The computing infrastructure is operated for the exclusive use of an organization. the cloud probably managed by the organization or third party. Private clouds are more costly and more secure when compared to public clouds. Private clouds may be either on or off premises. Externally hosted private clouds are also only used by one organization, but are hosted by third party specializing in cloud infrastructure. Externally hosted private clouds are inexpensive than On-premise private clouds.

2.2.3.Hybrid cloud combines multiple clouds (private community of public) where those clouds retain their unique identities, but are bound together as a unit. . A related term is Cloud Bursting. In Cloud bursting organization is use their own computing infrastructure for common usage, but access the cloud for high load requirements. This ensure that a sudden increase in computing necessity is handled gracefully. hybrid cloud may offer standardized or proprietary access to data and applications , as well as a application portability.

2.2.4.Community cloud a community cloud is one where the cloud has been organized to serve a common function or purpose. For example one organization or for several organization ,but they share conman concerns such as their mission, security, policies, regulatory compliance needs and so on.

3.CLASSIFICATION BASED UPON SERVICE PROVIDER

3.1 Infrastructure as a service (IaaS)

Infrastructure as a service (IaaS) involve offering hardware related services using the principles of cloud computing. IaaS provides a virtual-machine, virtual storage, disk image library, virtual infrastructure , raw block storage, and file or object storage, , load balancer, IP addresses, firewalls, virtual local area networks and software providers supply these resources on-demand from their large pools installed in data centers bundles. The IaaS service provider manages all the infrastructure. [10]

3.2 Platform as a Service (PaaS)

In the PaaS models, cloud providers deliver a computing platform, typically inclusive of programming language execution environment, operating system, database and web server. Application developers can evolve and run their software solutions on a cloud platform without the cost and insolubility of buying and managing the basic hardware and software layers. The service provider manages the cloud infrastructure, the operating systems and the enabling software.

3.3.Software as a service (SaaS)

includes a complete software offering at the cloud. Users can access a software application hosted by the cloud venders on pay-per-use basis. In the business model using software as a service , users are provided access to application software and databases. SaaS is a complete operating environment with applications , management, and the user interface .Cloud providers manage the infrastructure and platforms that run the applications. SaaS is sometimes designated to as "on-demand software" and is commonly priced on a pay-per-use backbone. SaaS providers generally price applications using a subscription fee. David Linthicum describes a more granular classification on the basis of service provided and listed there are below: [10]

- Storage-as-a-service
- Database-as-a-service
- Information-as-a-service
- Process-as-a-service
- Application-as-a-service

4. LOAD BANCING In CLOUD COMPUTING

Load balancing is the technique of distributing the load between various resources in any system. Thus load require to be distributed over the resources in cloud-based architecture, so that each resources 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. Cloud Load Balancers manage online traffic by distributing workloads between multiple servers and resources automatically. They maximize throughput, minimize response time, and avoid overload.

In this paper, an overall review of the latest load balancing technique in the Cloud Computing environment is submitted. The ideas of every algorithm are discussed and finally sum up as an overview. There are several issues while dealing with load balancing in a cloud computing environment. every load balancing algorithm must be such as to instate the required target.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 9, September 2014

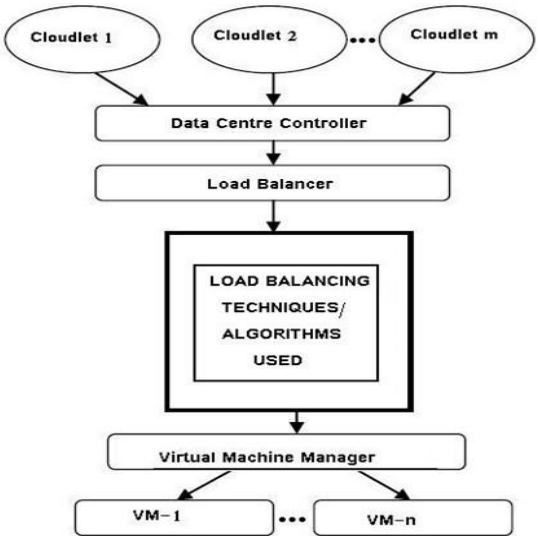

Fig 1: Load Balancing Algorithms Execution[10]

Some algorithms target at achieve higher throughput, some target at achieve minimum response time, some other target to achieve maximum resource utilization while some target at achieve a trade-off between all these metrics. Figure1 represent a framework underneath which various load balancing algorithms work in a cloud computing environment.

5. ISSUE EFFECTING LOAD BALANCING In CLOUD COMPUTING

There are many types of load balancing techniques which are available for cloud computing. These load balancing techniques are: geographical distribution, static and dynamic.[12]

5.1 The geographical distribution

The geographical distribution of the nodes matters a lot in the collective performance of any real time cloud computing systems, specifically in case of the big scaled applications like Twitter, Facebook etc. A well-distributed system of nodes in cloud environment is useful in handling fault tolerance and maintaining the efficiency of the system. Geographical load balancing (GLB) can be defined as a series of decisions about online assignment and/or migration of virtual machines (VMs) or computational tasks to geographically distributed datacenters in order to meet the service level agreements (SLAs) or service deadlines for VMs/tasks and to decrease the operational cost of the cloud system.[12]

5.2Static Load Balancing Algorithm[12]

In static load balancing algorithms, the execution of the processors is determined at the beginning of the execution, it does not depend on current state of the system. The aim of static load balancing is to decrease the overall execution time of a synchronous program while minimizing the communication delays.[12] These algorithms are mostly suitable for homogeneous and stable environments and can produce better results .Some of the examples of static load balancing algorithms are: Randomized algorithm, Round Robin algorithm and Threshold algorithm.

5.3 Dynamic Load Balancing Algorithm

In dynamic load balancing algorithm, the decisions in load balancing are based on the current state of the system, no prior knowledge is needed . The major advantage of dynamic load balancing is that if someone node fails, it will not stop the system, it will only affect the performance of the system.[12] These algorithms are more resilient than static algorithms, can easily adapt to alteration and provide better results in heterogeneous and dynamic environments. Dynamic load balancer uses morality for keeping the track of updated information. therein are four policies for dynamic load balancers: selection policy, transfer policy, location policy and information policy. The task of load balancing is shared among distributed nodes. In a distributed system, dynamic load balancing can be done in two different ways: distributed and non-distributed.

5.3.1 Distributed Dynamic Load Balancing Algorithm In the distributed one, the dynamic load balancing algorithm is executed by all nodes present in the system and the task of scheduling is shared among them. The interaction among the nodes to achieve load balancing can take two forms: cooperative and non-cooperative.[12]

5.3.2 Non-Distributed Load Balancing Algorithm

In the non-distributed or undistributed, the nodes work personal in order to instate a common goal. Non-distributed dynamic load balancing algorithms are ahead classified into two: centralized and semi-centralized.[12]

5.3.2.1 Semi-distributed Dynamic Load Balancing

In semi-distributed dynamic load balancing, the nodes of the system are divisions into clusters, where the load balancing in each cluster is of centralized form. A central node is elected in every cluster by appropriate election technique which takes care of load balancing within that cluster. Therefore, the load balancing of all system is done via the central nodes of each cluster. [12]

5.3.2.2 Centralized Dynamic Load Balancing

In centralized dynamic load balancing, the algorithm is only executed by a single node in the whole system i.e. central node. This node is perfectly responsible for load balancing of the whole system and rest of the nodes interacts only with the central node.[12]

The remaining part of this paper are organized as section 6 describes related work about various load balancing technique used in recently in soft computing , artificial antigens(AI) and other related subjects we conclude our paper and also provide direction for future enhancement.

6.RELATED WORK

This paper is designed for brief discussion about various Load Balancing techniques used in cloud architectures. In these techniques for Load Balancing of mentioning server load, main aims to guarantee that each computing resource is distributed effectively, ultimately to improve resource utilization. This paper is surveyed as:-

6.1 Stochastic Hill Climbing this paper [6][8], a soft computing based load balancing approach has been

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 9, September 2014

proposed. A local optimization approach Stochastic Hill climbing is used for allocation of incoming jobs to the servers or virtual machines (VMs). There are two main families of procedures for solving a optimization problem. Complete methods which guarantee either to find a valid assignment of values to variables or prove that no such assignment exists. These methods frequently exhibit good performance, and guarantee a correct and optimal answer for all inputs. Unfortunately, they require exponential time in the worst case, which is not acceptable in the cloud computing domain. The other incomplete methods may not guarantee correct answers for all inputs. Rather these methods find satisfying assignments for solvable problems with high probability.

A variant of Hill Climbing algorithm Stochastic Hill Climbing (SHC) is one of the incomplete approaches for solving such optimization problems. A stochastic and Local Optimization algorithm is simply a loop that continuously moves in the direction of increasing value, which is uphill. It stops when it reaches a peak where no neighbor has a higher value. This variant chooses at random from among the uphill moves and the probability of selection can vary with the steepness of the uphill move. Thus it maps assignments to a set of assignments by making minor changes to the original assignment. Each element of the set is evaluated according to some criteria designed to move closer to a valid assignment to improve the evaluation score of the state.

6.2 Particle Swarm Optimization (PSO) Algorithm

[2] Particle Swarm Optimization (PSO) as a meta-heuristics method is a self-adaptive global search based optimization technique introduced by Kennedy and Eberhart . The PSO algorithm is alike to other population-based algorithms like Genetic algorithms (GA) but, there is no direct recombination of individuals of the population. The PSO algorithm focuses on minimizing the total cost of computation of an application workflow.

As a measure of performance, Authors used cost for complete execution of application as a metric. The objective is to minimize the total cost of execution of application workflows on Cloud computing environments. Results show that PSO based task-resource mapping can achieve at least three times cost savings as compared to Best Resource Selection (BRS) based mapping for our application workflow. In addition, PSO balances the load on compute resources by distributing tasks to available resources.[1]

6.3 Genetic Algorithm

Genetic Algorithm (GA)[9] has been used as a soft computing approach, which uses the mechanism of natural selection strategy. A simple GA is composed of three operations: genetic operation, selection, and replacement. The advantage of this technique is that it can handle a vast search space applicable to complex objective function and can avoid being trapping into local optimal solution. A generation is a collection of artificial creatures (strings). [5]In every new generation, a set of strings is created using information from the previous ones. Occasionally, a new part is effort for good measure. GAs are randomized, but they are not simple random walks. They adept exploit historical information to speculate on new search points with expected improvement. The mechanics of a simple GA are amazingly simple, involving nothing much complex than copying strings and swapping partial strings. easier of operation and power of effect is two main attractions of the GA approach.[3] The effectiveness of the GA depends in appropriate mix of exploration and exploitation. Three operators to achieve this are: selection, crossover, and mutation But this technique is based on local search optimization technique and fetcher enhancement use in many global search optimization algorithm is use.

6.4 Ant Colony Optimization

they proposed an algorithm for load distribution of workloads among nodes of a cloud by the use of Ant Colony Optimization (ACO). This is a modified approach of ant colony optimization that has been applied from the perspective of cloud or grid network systems with the main aim of load balancing of nodes. This improving algorithm has an edge over the original approach in which each ant build their own individual result set and it is later on built into a complete solution. [5][11] However, in their approach the ants consecutive update a single result set rather the updating their own result set. Further, as they know that a cloud is the collection of many nodes, which can support various types of application that is used by the clients on a basis of pay per use. so, the system, which is incurring a cost for the user should function smoothly and should have algorithms that can continue the proper system functioning even at pinnacle us hours. ACO is inspired from the ant colonies that work together into foraging behavior. In fact the real ants have inspired several researchers for their work, and the ants approach has been used by many researchers for problem solving in different areas. [6]This approach is called on the name of its inspiration ACO.[1] The ants work totally in search of new sources of food and simultaneously use the existing food sources to shift the food back to the nest. The approach aims at efficiently distribution of the load among the nodes and such that the ants never encounter a dead end for movements to nodes for building an optimum solution set. In our algorithm, first a Regional load balancing node is chosen in CCSP, which will act as a head node.

6.5 Active Clustering

Active clustering is an enhanced method of random sampling, where this algorithm works on the principle of grouping similar nodes together and start working on these group nodes [12]. This method uses the resources efficiently thereby increases the throughput and performance of the system by using high resources. In this approach a technique called match-maker is introduced. When an execution starts in a network, the process gets and searches for the next matching node said to be match-maker which should satisfy the criteria that it should be the different one from the former one. Once the match-maker is found the process gets initiated and as soon as the process gets over the match-maker gets detached from the network. Thus this is an iterative process in the network to balance the load efficiently.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 9, September 2014

6.6 Honeybee Foraging Behaviour Algorithm

[6] Optimization algorithms are search methods where the goal is to find an optimal solution to a problem, in order to satisfy more objective functions, possibly subject to a collection of constraints. perusal of social animals and social insects have resulted in a number of computational models of swarm intelligence. in this paper [2] It is a decentralized honeybee-based load balancing technique that is a nature-inspired algorithm for self-organization . It achieves global load balancing via local server actions. Performance of the system is extended with increased system diversity but throughput is not increased with an increase in system size. It is best favorable for the conditions where the diverse population of service types is required.

6.7 Equally Spread Current Execution Algorithm

Equally spread current execution algorithm process handle with priorities. it distribute the load randomly by testing the size and transfer the load to that virtual machine which is lightly loaded or handle that task easy and take less time [4], and give maximize throughput. It is spread spectrum technique in which the load balancer spread the load of the job in hand into multiple virtual machines.

6.8 Throttled Load Balancing Algorithm [4]Throttled algorithm is completely based on virtual machine. In this client at first requesting the load balancer to check the right virtual machine which access that load easily and perform the operations which is give by the client or user. In this algorithm the client at first requests the load balancer to find a suitable Virtual Machine to perform the required operation.

Table 1. Comparison of existing Load Balancing Techniques

Algorithms	*Static environment*	*Dynamic environment*	*Centralized environment*	*Distributed environment*	*Hierarchical environment*
Round-robin	Yes	No	Yes	No	No
CLBDM	Yes	No	Yes	No	No
Ant Colony	No	Yes	No	Yes	No
Honeybee Foraging	NO	YES	No	Yes	NO
Particle Swarm Optimization	No	Yes	No	Yes	No
MaxMin	Yes	No	Yes	No	No
MinMin	Yes	No	Yes	No	No
Biased Random Sampling	No	Yes	No	Yes	No
Active Clustering	No	Yes	No	Yes	No
Stochastic Hill Climbing	No	Yes	No	No	Yes
OLB	Yes	No	Yes	No	No
Equally Spread Current Execution Algorithm	No	Yes	Yes	No	No
ESWLC	No	Yes	Yes	No	No
Genetic Algorithm	No	Yes	Yes	No	No

This metric is used to estimate the total number of tasks, whose execution has been completed successfully. High throughput is necessary for overall system performance.

7. CONCLUSION AND FUTURE WORK

This paper is based on cloud computing technology which has a very vast potential and is still unexplored. The capabilities of cloud computing are Interminable. Cloud computing provides everything to the user as a service which includes platform as a service, application as a service, infrastructure as a service. One of the major issues of cloud computing is load balancing because overloading of a system may lead to poor performance which can make the technology unsuccessful. So there is eternally a requirement of efficient load balancing algorithm for efficient utilization of resources. Our paper focuses on the different load balancing algorithms and their applicability in cloud computing environment.

In our future work we will discuss Simulated-Annealing for global search optimization in cloud load balancing and simulation. Some additional algorithms which can help in solving some sub-problems in load balancing which are applicable to cloud computing.

ISSN (Online) : 2278-1021
ISSN (Print) : 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 3, Issue 9, September 2014

REFERENCES

[1].Fazel Mohammadi , Dr. Shahram Jamali , and Masoud Bekravi "Survey on Job Scheduling algorithms in Cloud Computing"International Journal of Emerging Trends & Technology in Computer Science (IJETTCS) Volume 3, Issue 2, March – April 2014.

[2]. Baris Yuce , Michael S. Packianather ,Ernesto Mastrocinque , Duc Truong Pham and Alfredo Lambiase "Honey Bees Inspired Optimization Method: The Bees Algorithm" insects 1 July 2013; Published: 6 November 2013.

[3]. Bibhudatta Sahoo, Sudipta Mohapatra, and Sanjay Kumar Jena "A Genetic Algorithm Based Dynamic Load Balancing Scheme for Heterogeneous Distributed Systems" Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications, PDPTA 2008, Las Vegas, Nevada, USA, July 14-17, 2008,2 Volumes. CSREA Press 2008.

[4]. Ms.NITIKA "Comparative Analysis of Load Balancing Algorithms in Cloud Computing" International Journal of Engineering and Science Vol 1-Issue 1.

[5]. Kousik Dasguptaa, Brototi Mandalb, Paramartha Duttac, Jyotsna Kumar Mondald, Santanu Dame, A Genetic Algorithm (GA) based Load Balancing Strategy for Cloud Computing" International Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA) vol 10,2013.

[6]. Shilpa V Pius, Shilpa T S "Survey on Load Balancing in Cloud Computing" International Conference on Computing, Communication and Energy Systems (ICCCES-2014).

[7] . J. Uma, V. Ramasamy, A. Kaleeswaran " Load Balancing Algorithms in Cloud Computing Environment - A Methodical Comparison" International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 3, Issue 2, February 2014.

[8]. Brototi Mondala,∗ , Kousik Dasguptaa, Paramartha Duttab "Load Balancing in Cloud Computing using Stochastic Hill Climbing-A Soft Computing Approach" Procedia Technology 4 (2012) 783 – 789 2212-0173 © 2012 Published by Elsevier Ltd. doi: 10.1016/j.protcy.2012.05.128 C3IT-2012.

[9]. Albert Y. Zomaya, Senior Member, IEEE, and Yee-Hwei The "Observations on Using Genetic Algorithms for Dynamic Load-Balancing" IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 12, NO. 9, SEPTEMBER 2001.

[10]. N. S. Raghava and Deepti Singh "Comparative Study on Load Balancing Techniques in Cloud Computing" OPEN JOURNAL OF MOBILE COMPUTING AND CLOUD COMPUTING Volume 1, Number 1, August 2014.

[11]. Bharti Suri "Implementing Ant Colony Optimization for Test Case Selection and Prioritization" Bharti Suri et al. / International Journal on Computer Science and Engineering (IJCSE).

[12]. Shiny "Load Balancing In Cloud Computing:A Review" IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 15, Issue 2 (Nov. - Dec. 2013), PP 22-29.

[13]. Mohiuddin Ahmed, Abu Sina Md. Raju Chowdhury, Mustaq Ahmed, Md. Mahmudul Hasan Rafee "An Advanced Survey on Cloud Computing and State-of-the-art Research Issues" IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 1, January 2012.