Comparative Load Balancing Algorithms in Cloud Computing

Mohammadjavad Hosseinpoor*, Fatemeh Rezvani, Somayeh Sajadi,

***Abstract*— Cloud computing as a model based on wide computer networks such as the Internet which is the new model for supply, utilization, delivery of IT services and other shared resources by using the Internet to provide computing. Recently, from starting point of internet task until now, has an effective a lot changes that some of them Cause Changes the method of human Life at Several Decade. One of the most modern changes at How does the internet work, is that by Introducing cloud computing. In order to its characteristics of these modern technology has been quickly popular which is why in cloud computing all type of Facilities delivery to users as a service. Naturally, Each Change and the new concept in technology world, has a special difficulties and complexity. Using from cloud computing has many challenges that faced by experts in this field that include: Load balancing, security, reliability, ownership, data backup and portability. According to importance of load balancing process in cloud computing, the goal of the present study is to compare the process and techniques in this field. Grid computing is a new model to develop heavy parallel computers that their goal is to create abstract super computer with use free source in WAN networks for unreliable platforms such as internet. Since in Grid computing platform communication and source entity is so dynamic, also resource management including load balancing is an important issue. First, in this paper we describe and define about cloud computing, advantages and disadvantages and then algorithms compare in this filed.**

***Index Terms*— Cloud computing, load balancing, heterogeneous environments.**

I. INTRODUCTION

Cloud computing is a type of Internet-based computing that provides shared computer processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources (e.g., computer networks, servers, storage, applications and services), which can be rapidly provisioned and released with minimal management effort. Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in either privately owned, or third-party data centers that may be located far from the user–ranging in distance from across a city to across the world[5]. Cloud computing relies on sharing of resources to achieve coherence and economy of scale, similar to a utility (like the electricity grid) over an electricity network. Advocates claim that cloud computing allows companies to avoid up-front infrastructure costs (e.g., purchasing servers). As well, it enables organizations to focus on their core businesses instead of spending time and money on computer infrastructure. According to the official NIST definition:

cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction[3].

Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in either privately owned, or third-party data centers that may be located far from the user–ranging in distance from across a city to across the world [12-13]. Many of cloud computing services was delivered, with using utility computing model enable to provide these services as a similar way. However other type of applies based on its own shared services. Sharing computing power among multitenancy has been unable to improve productivity rate, because with this way servers never disable without any reason. One way that can mentioned is computers used as a high amount in a way that client's cloud computing don't need to calculate and determine Peak Load.

II. LOAD BALANCING

Cloud computing one of the modern changes in IT (information technology) and gradually have been extended. From participate in market view, understanding effectiveness about load balancing is an important in cloud computing. Cloud computing platform as an automatic platform as a service that allow users to:

- buy
- Remote response
- Dynamic Scalability and
- Manage system which is currently load balancing in cloud computing system as a challenges.

Always distributed solution is needed, because maintained one or more disable server is impossible to do on demanded needs or is not useful. It's clear that in order to scale and

Manuscript received Feb, 2017.

**Mohammadjavad Hosseinpoor, Member of faculty in Department of Computer Engineering and Member of Young researchers and elite club, Estahban Branch , Islamic Azad University, Estahban,Iran.*

***Sara Rezvani**, Department of Computer Engineering and Member of Young researchers and elite club, Estahban Branch , Islamic Azad University, Estahban,Iran.*

***Somayeh Sajadi**, Department of Computer Engineering and Member of Young researchers and elite club, Estahban Branch , Islamic Azad University, Estahban,Iran.*

complexity of these systems, it's impossible to focus on the other servers.

We need to load balancing to manage appropriate these service provider resource which is delivery to offer service. Assessment of level load usually was Mechanization to continuous one service while one or more set of component is going to be failed.
Components are regularly Monitoring and when one of the components aren't response, load balancing goes up and the traffic is sent. With Assessment of level load at utilization area often can peak load problem which is not only decrees cost and create Green computing, but the pressure on individual circuits keep in low level which is increase potentially their life time. In fact we can say, the aim of load balancing is to find appropriate Mapping from current processors on systems task in a way that each processor, running equal amount of task until total running time decrees to lowest amount [4].
Load balancing is enable balancing to load with dynamic sending local capacity of task from a machine to another machine on remote node and / or the machines which is less usable [19-20]. This operate of user satisfies at least: mechanization of source will be increase, number of task will be decrease and efficiency of system will be increase. Also, to access green computing in cloud, the load balancing will be needed.
The responsible factors has mentioned below:
Limitation of energy consumption: the load balancing can avoid from too Interaction nodes or virtual machines decrease amount of energy consumption according to capacity of task.
Reduce carbon emissions: Energy consumption and carbon emissions are two sides of the same coin. Both are directly proportional to each other. Load balancing helps to reduce energy consumption that are automatically leads to a reduction in carbon emissions, and therefore comes to green computing.
The goal of load balancing:

- Significant improvements in an efficiency
- Having a program and backup plan while the system or even part of it is fractured.
- To maintain system stability
- For future reforms in the system

A. Review Stage

Submit your manuscript electronically for review.

B. Final Stage

When you submit your final version, after your paper has been accepted, prepare it in two-column format, including figures and tables.

C. Figures

As said, to insert images in *Word*, position the cursor at the insertion point and either use Insert | Picture | From File or copy the image to the Windows clipboard and then Edit | Paste Special | Picture (with "Float over text" unchecked).

The authors of the accepted manuscripts will be given a copyright form and the form should accompany your final submission.

III. LOAD BALANCING ALGORITHMS

Depending on who start process of load balancing, algorithms can be classified in three categories:

- Sender Initiated: If the load balancing algorithm is started by the Sender.
- Receiver Initiated: If the load balancing algorithm is started by the receiver.
- Symmetric: a combination of both method that is mentioned above

According to system current mode, the load balancing algorithms can be classify in two categories:

- Static: it isn't referee to system current status, and need for prior knowledge of the system.
- Dynamic: compared to the system current status and decisions on the load balancing.

A. static load balancing

Static load balancing, allocates a task to a processor or fix node. When the system restart, the same processors binding to consider without any changing that may be use in during the system life time occurred. May be it doesn't have result to the same processor.

But the task is fix in arrange or model. The Efficiency of virtual machines will determine the time of task achievement.

Master processor is allocated Workload to other Slave processors according to their efficiency. Dedicated task has been performed by processor and the result is returned to the Master processor [8].

Static load balancing algorithms allocated tasks to the node is only based on their ability to process. This process is just depend on previous knowledge about node which is contain:

- Node Power processing
- Memory
- Storage capacity and it's known the latest efficiency knowledge of communications

Although may include previous efficiency knowledge communications. Static algorithms often aren't consider running dynamic changes. In addition, these type of algorithm couldn't consistency with changing load during running and their goal is to minimize the time of running task is limitation Overload communication and postpone [2].

B. Dynamic load balancing

In Dynamic load balancing algorithms, workload distributed at run-time among processors. Master processor, allocated the new collective information to the slave processor. Dynamic load balancing algorithms accept:

- Properties of nodes,
- Capabilities and
- Network bandwidth.

The current algorithms insist on combination of data about nodes which was collective in cloud and properties about their run-time.

The algorithm allocates tasks and may be according to

dynamic and return it to the nodes with collective and calculated properties, it needs to be continual monitoring, task flow and often its complexity to implementation. But it has a high accuracy and can produce optimize load balancing outcome.

One of the advantages about dynamic load balancing is that load balancing decisions according to current state system which is to help total improvement efficiency with departure dynamic load. A dynamic strategies sometimes running several times and it is possible to re-schedule a task to a new node according to allocate dynamic system environment. [9]

Dynamic load balancing can be done in two different ways:

- Distributed
- Non-distributed

In distributed, dynamic load balancing algorithms implemented by all nodes in the system and load balancing task is divided among all of them. The Interaction between nodes to achieve load balancing have in two forms:

- Collaboration
- Non- collaboration [5].

The first form of nodes work together to achieve a common goal. In the second form each node work is independent in a directed local goal because the nodes within a system needs to communicate with any other node. The advantage of this method is that if one or several nodes is fail, the load balancing wouldn't make the total of load balancing process to stopped, but somewhat effect of system performance. [12]

Dynamic load balancing can be distributed on a system where each node to exchange status information with other node within system needs, creates a lot of pressure. This approach is more beneficial than individual nodes that has little interaction with other nodes. [2]

- In an undistributed type, a node or group of nodes has to do the task of load balancing. Dynamic load balancing algorithms have been focused on both centralized and Semi-Distributed.
- In a centralized type, the load balancing algorithm implemented only a system node (central node). These nodes alone responses to the task of total load balancing system. The central nodes only interact with other nodes. In semi-distributed system nodes divided into clusters where each cluster of load balancing is centralized based.

With Select central node to hold on load balancing in cluster by suitable techniques for each cluster. The overall system load balancing is performed by nodes centralized dynamic load in the center of each cluster, less messages will receive.

Therefore, the total number of interactions within a system is reduced to a Semi-Distributed. However, centralized algorithms can create a central bottleneck node and when central node balance falls there isn't any results. Therefore, this method is suitable for small-sized networks [2].

.

IV. COMPARE LOAD BALANCING ALGORITHMS

In this section, various load balancing algorithms expressed in cloud computing environments, as well as The algorithm based on static or dynamic environment and the balance distributed or centralized or hierarchical which they have been compared with each other.

A. Active clustering algorithm

Grouping nodes have a similarly task with other nodes. The process includes: a node start process and from between valid own neighbor nodes select an Interface node rather than pervious node. Then Interface node establishes a connection between a neighboring nodes as type of primary forms. Then interface node separated primary node with own node. The set of processing that is mentioned above regularly repeated. In this algorithm, with increasing resource, efficiency of system will increase. As a result, by using efficient source throughput will be high. With increasing system diversity the current algorithm is being reduced.

B. Bees search algorithm

A state of demand is increasing or decreasing in web server load balancing, allocation services to dynamically adjust user demands. The servers as virtual servers are grouped and each virtual server has its own virtual service queues. Similar boards here are the bee dance. Also this page to advertise the benefits (required interest) are used in total colonies. Each of these servers as a role of a browser or a watch [12]. The benefits of server processing is that request can be sent to the billboards with rp probability. One server can select a queue of virtual servers with the probability xp (represents the behavior of the browser) or can examine ads (see the dance) and service that it showed scouts behavior. A server is request a service, after calculating and comparing its interests with the common interests of the colon, the amount of xp is consideration. If this interests remains is high, the current virtual server on the same server to send an rp probability, if the interest is low, the server load balancing method as a scout or general probe discovers through the activities of local servers. With increasing diverse system, the efficiency of system will be increased. But throughput with increasing system, won't be increase [7].

C. Random sampling orientation

Here is a virtual graph where each node (server as a node) the server once a connection to the display. In graph, each server as a node with any degree of free resources directed. While a node do a task, it's deleted from input edge which is indicated decrease access to free resource. After completing a task, node create one input edge which is indicated increasing access to free source. Adding and removing process is done by Random sampling. At the beginning path of each node randomly selected a neighbor. Last node for load allocating be selected. In addition, other methods can be used to select a node for load allocation is based on specific criteria. However, the selection and allocation load can be used to nodes with low load. Finally, what we going to be taken is directed graph, this method balances is to load across all system nodes. Sources increases similar population and

high efficiency of the system, thus increasing the throughput of the system resources increases load-balancing design that is totally decentralized, so using that large networks such as cloud systems are appropriate [14].

D. Cardboard method

To control the use of distributed load balancing and rate limiting (DRL), Cardboard mechanisms has proposed for Cloud Services. With computational and communication overhead is very low, this algorithm is simple and easy to implement. The environment of using this algorithm, unified framework to control cloud. From among load balancing criteria, the criteria has taken into overhead and resources.

E. Event-based method:

A load balancing algorithm event - based has suggested for online multi-player gaming massive real time (MMOG). This algorithm after receiving capacity as incoming events, its components and resources is analyzed in the field of general Game Session. As a result, load balancing activities, produced Game Session. This method is able to scale a Game Session based on multiple sources of user variable load increase or decrease But sometimes service quality and non-infringement. In this algorithm, sometimes through load balancing criteria, the standard is intended utilization of resources.

F. Load balancing server- based for distributed Web services

A policy load balancing service-based for Web servers which has been proposed distributed worldwide. The policy to reduce the time limit number of changes service using routes a call to the nearest remote server helps without overloading them. A middleware to implement this protocol has been described, also uses a heuristic method to help Web servers to handle the load tolerance. [18]

G. Fuzzy logic:

A load balancing algorithm based on turn circulating in virtual machine environment in cloud computing has designed in order to access better response and processing time. Load balancing algorithm has been done before Process Servers come.

The task is scheduled on the basis of various parameters like:

- CPU speed,
- Time allocated to virtual machines, etc.
- The present algorithm storage information on each:
- virtual machine and
- The number of requests which is currently allocated to the virtual machine.

Therefore Fuzzy logic is look like natural language, it can formulate and edit their own problems. In this architecture Fuzzier runs Fuzzification process is converts two types of input data such as processor speed and the load allocated to a virtual machine and an output such as load balancing that is required in inference system. The plan of CPU speed and load virtual machine as a two input parameters, in order to consider better load balancing in cloud which is using fuzzy logic. These parameters are given as input to Fuzzier to measure load balancing that used as output[15-16]. Two parameter called as a processor speed and load-allocated virtual machine that is to assess load-balanced on data centers in cloud computing environments are used through fuzzy logic. The speed of Processor and time allocated to the virtual machine applied through fuzzy logic to load balancing in cloud computing [11].

H. Message-oriented model

Clusters provide to take advantage of distributed applications by different computers on networks. The current issue which is related to cluster Bring up on network performance, if the total load is distributed by a computer network distribution, cause slowing task on network[10]. To prevent this situation, Resource Management Software Metrics can be used for traffic distribution between sites, so that network performance is maintained at a high level.

Web services are mainly used offline instant messaging applications, this technology is for real-time communication between different partners, and however the presence and availability of applications is important. A model was introduced to uses the XMPP for load balancing. XMPP clients send current information to current server XMPP and flows XML detailed current information which is generated by these servers have clients in practice. Using a load balancing above XMPP server allows incoming requests by public services, priorities and criteria to be employed the load balancing of algorithm from among load balancing consider response time and efficiency [6].

I. Min-Min algorithm

This algorithm started has not been allocated a series of tasks, First of all, the minimum completion time for all tasks can be found. Then between minimum times, the lowest choose the minimum times of all tasks on whatever source is available. Then, according to the minimum times, the task on the machine is scheduled. Then runtime for all other tasks on the machine by adding the time of the assigned task the execution time on other tasks upgraded machine. It is allocated from the list of tasks to devote be deleted.

The same procedure again until all tasks assigned to the sources, follows. But this method has a major drawback that can lead to hunger. [17] In this algorithm balances the load among different criteria, measure resource utilization, overhead, throughput, response time and efficiency is considered [1].

J. Min-Max algorithm

This algorithm is almost the same as Min-Min algorithm except the following factors: After finding the execution times at least, the highest value is selected the maximum time between all duties on any available source. Then, according to the maximum time, the task is scheduled on the machine.

K. The two-stage load balancing algorithms OLB + LBMM

A two-stage scheduling algorithm is proposed that

scheduling algorithms OLB (opportunities load balancing) and load balancing LBMM, Min-Min combined for the use of improved performance and balance. OLB scheduling algorithm, each node to achieve the purpose of load balancing, scheduling algorithm LBMM mode hold task to reduce the execution time of each task on node, thus reducing the total execution time. Explore the bees, provide random sampling and active clustering. [1]

L. Queue-Idle-Join algorithm

A load balancing algorithm suggests for web services, dynamic scalability. The load balancing algorithm in large-scale distribution provides with distributors.

Using this algorithm environment, data centers and the cloud effectiveness criteria, response time and overhead. [7]

M. The central load balancing policy for virtual machines

A central load balancing policy for virtual machines have suggested that the load equally in a cloud computing environment or virtual machine distributed balances. This policy increases overall system performance, but systems are not considered with fault tolerance.

This general mode of information for decisions balance increases more than 20% efficiency. Using this algorithm cloud computing environment is intended the effectiveness criteria, response time, throughput and resource utilization. [7]

V. CONCLUSION

One of the main challenges in cloud computing is load balancing evenly distributed workload across all nodes and dynamic local to achieve high user satisfaction and utilization rate of resources by ensuring the allocation of compute resources is equitable and efficient.

In this paper, we compared different algorithms and load balancing in cloud computing which is conclude that we can use a special algorithm according to our needs. Therefore, develop of matching algorithm is that suitable for heterogeneous environments and reduce costs.

REFERENCES

[1] Sran.N and Kaur.N, Comparative Analysis of Existing Load Balancing Techniques in Cloud Computing, International Journal of Engineering Science Invention, Vol.2, Issue 1,2013.
[2] Khetan.A, Bhushan.V and Chand Gupta.S, A Novel Survey on Load Balancing in Cloud Computing International Journal of Engineering Research & Technology (IJERT) , Vol.2, Issue 2 ,2013.
[3] Ranjith Kumar.Y, Madhu Priya.M and Shahu Chatrapati.K, Effective Distributed Dynamic Load Balancing For The Clouds, International Journal of Engineering Research & Technology (IJERT), Vol.2, Issue 2, 2013.
[4] Begum.S and Prashanth.C.S.R, Review of Load Balancing in Cloud Computing. IJCSI International Journal of Computer Science Issues, Vol.10, Issue 1, 2013.
[5] Mohammadjavad Hosseinpour**,** Mortaza Zolfpour, 2015, "An investigation in cloud computing security: problems and challenges", International Journal of Scientific and Engineering Research.
[6] Hamo.A and Saeed.A, Towards a Reference Model for Surveying a Load Balancing, IJCSNS International Journal of Computer Science and Network Security, Vol.13, 2013.
[7] Rastogi.D, Bansal.A and Hasteer.N, A Techniques of Load Balancing in Cloud Computing: A Survey, International Conference on Computer Science and Engineering (CSE) 7th, April 2013.
[8] Kansal.N.J and Chana.I, Cloud Load Balancing Techniques: A Step towards Green Computing, IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 1, January 2012.
[9] Rajguru.A and A.S, A Comparative Performance Analysis of Load Balancing Algorithms in Distributed System using Qualitative Parameters, International Journal of Recent Technology and Engineering (IJRTE), Vol.1, Issue 3, 2012.
[10] Al Nuaimi.K, Mohamed.N, Al Nuaimi.M and Al-Jaroodi.J, A Survey of Load Balancing in Cloud Computing: Challenes and Algorithmsg, 2012 IEEE Second Symposium on Network Cloud Computing and Applications, 2012.
[11] Membrey.P, Hows.D and Plugge.E, Load Balancing in the Cloud, pp.211 – 224, 2012.
[12] Sethi.S, Sahu.A and Jena.S.K, Efficient load Balancing in Cloud Computing using Fuzzy Logic, IOSR Journal of Engineering (IOSRJEN), Vol. 2, Issue 7, pp.65-71, 2012.
[13] P. Mell and T. Grance. (2009, Oct.). The NIST definition of cloud com-puting, NIST special publication 800-145, National Institute of Standards and Technology. Available: http://www.csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf.
[14] S. Marston, et al., "Cloud computing — the business perspective," Decision Support Systems, vol. 51, pp. 176-189, 2011.
[15] S.Iqbal,, G.F.Care, "Performance Analysis of Dynamic Load Balancing Algorithms with Variable Number of Processors", Journal of Parallel and Distributed Computing, Vol. 65, pp. 934-948, 2005.
[16] S. RAOGURURAJ, S. STONEHAROLD and T.C. HU, "Assignment of Tasks in A Distributed Processor System with Limited Memory", IEEE Trans. On Computers, Vol. C- 28, No. 4, April 1979.
[17] Comparative Study on Load Balancing Techniques in Cloud Computing N. S. Raghava* and Deepti Singh
[18] Chiang.M.L, Luo.J.N, Lin.C.B and Wang.S.S, High-Reliable Dispatching Mechanisms for Tasks in Cloud Computing. Department of Information and Communication Engineering, Chaoyang University of Technology, 2011.
[19] Hung.C, Wang.H and Hu.Y, Efficient Load Balancing Algorithm for Cloud Computing Network, 2011.
[20] Mondal.B and D.K, Load Balancing in Cloud Computing using Stochastic Hill Climbing-A Soft Computing Approach, Published by Elsevier Ltd, 2011
[21] M. Dorigo and L. M. Gambardella, "Ant colony system: a cooperative learning approach to the traveling salesman problem," IEEE Transactions on Evolutionary Computation, Volume 1, No.1, 1997, pp. 53-66.