Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 11, November 2014, pg.175 – 180

RESEARCH ARTICLE

Best Partition Searching In Public Cloud

Siddani Murali[1], B.JayaLakshmi Reddy[2]

PG Scholar in the Department of Computer Science and Engineering TKR College of Engineering and Technology, Hyderabad, Telangana

Assistant Professor Department of Computer Science and Engineering TKR College of Engineering and Technology, Hyderabad, Telangana

Abstract: Now a days cloud computing is one of the extreme platform which provides storage of data in very lower cost and available for all time over the internet. But it has more critical issue like security, load management and fault tolerance. Here we are discussing Load Balancing approach. Many types of load concern with cloud like memory load, CPU load and network load. Load balancing is the process of dispensing load over the different nodes which provides good resource exploitation when nodes are overloaded with job. Load balancing has to handle the load when one node is overloaded. When node is overloaded at that time loads scattered over the other ideal nodes. Many algorithms are available for load balancing like Static load balancing and Dynamic load balancing. Load balancing in the cloud computing location has an important impression on the performance. Good load balancing makes cloud computing more efficient and improves user satisfaction. A better load balance model for the public cloud based on the cloud segregating concept with a switch mechanism to choose different strategies for different situations. The algorithm applies the game theory to the load balancing strategy to improve the efficiency in the public cloud atmosphere

Keywords: Cloud Computing, Load balancing, Virtualization, load balancing model, public cloud, cloud divider, game theory

I. Introduction

The term cloud initiates from the world of broadcastings when providers began using virtual private network provisions for data announcements. The explanation of cloud computing provided by National Institute of Standards and Technology (NIST) says that: Cloud computing is a model for enabling convenient, on demand network access to a Public pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or provision provider interaction. So through this cloud computing there is no need to store the data on desktops, portables etc. You can store the data on servers and you can access the data through internet. Cloud computing provides better utilization of scattered resources over a large data and they can access remotely through the internet.

II. Load Balancing

Load Balancing is a technique in which the workload on the resources of a node is shifts to individual Resources on the other node in a network without troubling the running task. A standard way to rule web applications is by using a hardware based load balancer. The load balancer assumes the IP address of the web application, so all announcement with the web application hits the load balancer first. The load balancer is connected to one or more identical web servers in the back end. Depending on the user session and the load on each web server, the load balancer forwards packets to different web servers for processing. The hardware based load balancer is designed to handle high-level of load, so it can easily scale .However, a hardware based load balancer uses application specific hardware-based components, and thus it is typically expensive. Because of cloud's commodity business model, a hardware-based load balancer is rarely occurred by cloud providers as a Provision. Instead, one has to use a software based load balancer running on a generic server

The goals of load balancing are:

- To improve the performance substantially.
- To have a backup plan in case the system fails even partially.
- To maintain the system stability.
- To accommodate future modification in the system.

Types of Load Balancing Algorithms

1) Static Algorithms

Static algorithms divide the traffic equivalently between servers. By this approach the traffic on the servers will be disdained easily and consequently it will make the situation more perfectly. This algorithm, which divides the traffic equally, is announced as round robin algorithm. However, there were lots of problems appeared in this algorithm. Therefore, weighted round robin was defined to improve the critical challenges associated with round robin. In this algorithm each servers have been assigned a weight and according to the highest weight they received more connections. In the situation that all the weights are equal, servers will receive balanced traffic

2) Dynamic Algorithms

Dynamic algorithms designated proper weights on servers and by searching in whole network a lightest server Preferred to balance the traffic. However, selecting an appropriate server needed Real time

III. Prevailing Method

In single storage cloud system each cloud customer's data is stored on single higher configuration server. Even if that server has huge amount resources such as RAM, Hard disk, processing power, it has certain limit. If it crosses that limit then particular resource performance slows down.

IV. Projected Method

The load balancing model given in this article is aimed at the public cloud which has numerous nodes with scattered computing resources in many different physical locations. Thus, this model divides the public cloud into several cloud dividers. When the atmosphere is very large and complex, these divisions simplify the load balancing. The cloud has a main controller that chooses the suitable dividers for arriving jobs while the balancer for each cloud divider chooses the best load balancing strategy.

The load balancing strategy is based on the cloud segregating concept. After creating the cloud dividers, the load balancing then starts. When a job arrives at the system, with the main controller deciding which cloud divider should receive the job. The divider load balancer then decides how to assign the jobs to the nodes. When the load status of a cloud divider is normal, this segregating can be accomplished locally. If the cloud divider load status is not normal, this job should be transferred to another divider.

4.1 Projected Method Architecture

There are several cloud computing categories with this work focused on a public cloud. A public cloud is based on the standard cloud computing model, with provision provided by a provision provider. A large public cloud will include many nodes and the nodes in different physically locations. Cloud segregating is used to manage this large cloud. A cloud divider is a subarea of the public cloud with divisions based on the physical locations. The architecture is shown in Fig.1. The load balancing strategy is based on the cloud segregating concept. After creating the cloud dividers, the load balancing then starts: when a job arrives at the system, with the main supervisor deciding which cloud divider should receive the job. The divider load balancer then decides how to assign the jobs to the nodes.

Main Supervisor and Balancer

The load balance solution is done by the main controller and the balancers. The main controller first assigns jobs to the suitable cloud divider and then transfers with the balancers in each divider to refresh this status information. Since the main controller deals with information for each divider, smaller data sets will lead to the higher processing rates. The balancers in each divider gather the status information from every node and then choose the right strategy to distribute the jobs. The relationship between the balancers and the main controller is shown in Fig.1.

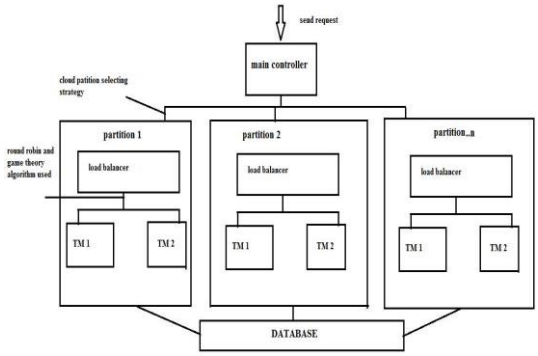

Fig.1 Relationship between Balancer and Main Controller

Assigning Jobs to the Cloud Divider

When a job arrives at the public cloud, the first step is choose the right divider. The cloud divider status can be divided into three types:

(1) Idle: When the percentage of idle nodes exceeds α, change to idle status.

(2)Normal: When the percentage of the normal nodes Exceeds β, change to normal load status.

(3) Overload: When the percentage of the overloaded nodes exceeds γ, change to overloaded status.

The parameters α, β, and γ are set by the cloud Divider balancers.

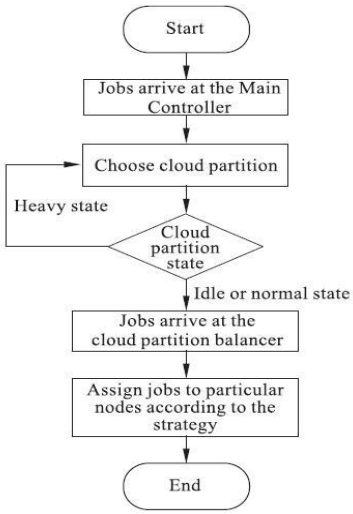

Fig.2 Flowchart for Job Assignment

Functioning Atmosphere

Operational atmosphere means atmosphere in which user interact with the application. For example, the DOS atmosphere consists of all the DOS commands available to users the windows atmosphere, on the other hand, is a graphical user interface uses icons and menus instead of commands. In proposed system user interacts with the application in following atmosphere.

1) A Personal Computer.
2) Windows Operating System.

Techniques Used
Our system is based on the following techniques:

Divider Selection Technique

Q= {A, B, C}

Γ={0,1}
b= -1 (“
q_0=A (initial
State) F=C
Process:
Turing Machine1: Calculating load degree
1. Compute Load $\sum$

Degree Inputs:
The static parameters include the number of CPU’s, the CPU processing speeds, the memory size, etc. Dynamic parameters are the memory utilization ratio, the CPU utilization ratio, the network bandwidth.

Process:-
1. Define a load parameter set: F= {F1,F2…Fm} with each Firepresents the total number of the parameters.
2. Compute the load degree as
Load Degree(N)= ∑ αiFi
Where i= 1…m
3. Average cloud divider degree from the node load degree statistics as:
Load degreeavg=∑i=1..n LoadDegree(Ni)

4. Three level node status are defined
Load_degree(N)=0 for***Idle***
0<Load_Degree(N)<Load_Degree(N)high for ***Normal***
Load_Degree(N)high <= Load_Degree(N) for ***Overloaded***
Output :-
Idle or Normal Or Overloaded

2. No N Cooperative load balancing game Input:-
Sji be the fraction of jobs that user j send to computer i
The vector sji =(Sj1,Sj2……Sjn) is called the *load balancing strategy* of user j.
The vector Sj = (S1,S2, ….Sm) is called the *strategy profile* of the load balancing game

Process
1. The expected response time at computer I
is Fi(S) = 1/ µi-∑ j=1 to m sjiϕk
2. The overall expected response time of user j is given by
Dj(S)= ∑i= 1 to n sji Fi(S) = ∑i=1 to n sji/µi-∑k=1 To m skiϕk
3. The goal of user j is to find a feasible load balancing strategy S ji such that Dj(S)is minimized.

Output:-
The decision of user j depends on the load balancing decisions of other users since Dj(S) is a function of S

Turing Machine 2:Get Avg. Load Degree & Classify Load

Q={ A,B,C,D,E } i.e. C= idle, D= Normal, E= Overloaded
Γ={0,1}
b= -1 (“
q_0=A (initial
State) F=C
a) Get Avg. Load Degree
$\text{Load\_degree}_{avg} = \sum$
b) Classify load
Idle
Load degree (n) =0

Normal

$0 < \text{load degree}(n) <= \text{load\_degree}_{high}$

Overload

$\text{load\_degree}_{high} <= \text{load degree}(n)$

The above Turing machines works for calculating load degree & classifies loads in three states i.e. idle, normal, and overloaded. It just makes the summation of all static & dynamic weights of processes. Since machine has single Finite State & No other halting condition this problem comes under NP type problem. Also, It gives result in polynomial time, So, it is P type.

Round Robin Scheduling

Assignment of exact resource to incoming request is **NP Hard** Problem. Round Robin Scheduling Algorithm makes slices of request to minimize overloads.

We assign the computing resource based on their load degrees, So problem gets reduced to **NP-Complete**.

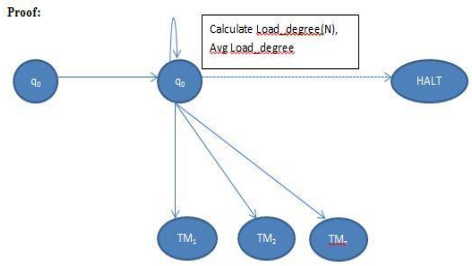

Fig.3 Processing through Turing Machine

- To improve the performance substantially
- To have a backup plan in case the system fails even partially
- To maintain the system stability
- To accommodate future modification in the system

Conclusion

Cloud computing system has widely been adopted by the industry though there are many existing issues like load balancing, migration of virtual machine, server unification which have been not yet fully addressed. Load balancing is the most central issue in the system to distribute load in efficient manner. It also ensures that every computing resource is scattered efficiently and fairly. Existing load balancing technique have been studied mainly focus on reducing overhead, reducing migration time and improving performance.[

References

[1] Gaochao Xu, Junjie Pang, and Xuedong Fu"**A Load Balancing Model Based on Cloud Segregatingfor the Public Cloud**" IEEE TRANSACTIONS ON CLOUD COMPUTING YEAR 2013

[2] K. Nishant, P. Sharma, V. Krishna, C. Gupta, K. P. Singh,N. Nitin, and R. Rastogi, **Load balancing of nodesin cloud using ant colony optimization, in Proc**. 14thInternational Conference on Computer Modelling andSimulation (UKSim), Cambridgeshire, United Kingdom,Mar. 2012, pp. 28-30

[3] M. Randles, D. Lamb, and A. Taleb-Bendiab, **Acomparative study into scattered load balancingalgorithms for cloud computing,** in Proc. IEEE 24thInternational Conference on Advanced InformationNetworking and Applications, Perth, Australia, 2010,pp. 551-556

[4] Ms. Parin V. Patel, Mr. Hitesh. D. Patel, Pinal. J. Patel,**A Survey On Load Balancing In Cloud Computing**, International Journal of Engineering Research & Technology (IJERT)Vol. 1 Issue 9, November- 2012ISSN: 2278-0181

[5] MicrosoftAcademicResearch,Cloud putting,"**http://libra.msra.cn/Keyword/6051/cloud-computing?query=cloud%20 "**computing, 2012

[6] R. Hunter, The why of cloud, **http://www.gartner.com/DisplayDocument?doc cd=226469&ref**= g noreg, 2012.

[7] N. G. Shivaratri P. Krueger, and M. **Singhal, Loaddispensing for locally scattered systems, Computer,vol. 25, no. 12, pp. 33-44, Dec. 1992**. B. Adler, Load balancing in the cloud: Tools, tips and

[8] Techniques **http://www.rightscale. com/infocenter/whitepapers/Load-Balancing-in-the-Cloud.pdf,2012**

[9] D. MacVittie, Intro to load balancing for developers The algorithms, **https://devcentral.f5.com/blogs/us/introto-load-balancing-for-developers-ndash-the-algorithms**,2012

[10] Anthony T.Velte, Toby J.Velte, Robert Elsenpeter, **Cloud Computing A Practical Approach** TATA McGRAW-HILL Edition 2010.

[11] D. Grosu, A. T. Chronopoulos, and M. Y. Leung**, Loadbalancing in scattered systemsAn approach usingcooperative games, in Proc**. 16th IEEE Intl. Parallel andScattered Processing Symp., Florida, USA, Apr. 2002, PP. 52-61.

Authors

1) S.Murali PG Scholar in the Department of Computer Science and Engineering TKR College of Engineering and Technology. Hyderabad, Telangana.

2) B. JayaLakshmi Reddy Assistant Professor Department of Computer Science and Engineering TKR College of Engineering and Technology. Hyderabad, Telangana

