International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 09 | Sep 2018 www.irjet.net p-ISSN: 2395-0072

A Novel Load Sharing technique for Server Farm Process Optimization

M.Ramya[1], Bh.V.RamaKrishna[2], S.Guna[3], B.SaiSunayana[4], P.KarishmaRani[5]

[1,3,4,5]B.Tech. (CSE)-Student, ISTS-Women's Engineering College, A.P, INDIA
[2]Research Wing Member, Acharya Nagarjuna University, Guntur, A.P, INDIA

Abstract – *The scope of dynamic load balancing mechanisms in Cloud Server technology increased due to high dependency of clients through heterogeneous environments. In this paper an attempt made to provide a novel technique to improve the percentage of high scale processor availability to support big process requests decreasing the rate of server busy response. This approach can be adopted to tune up load balancing over server farms used by Cloud Servers in SAAS / PAAS / IAAS services. The file size downloaded by client considered in this paper. Also cloud server metrics can be used to calculate a Load-Score (Ω) by which the client processes are ordered using a data structure. Three varieties of schemes are used to apply load sharing among Server-Farms. A software module developed to perform analysis over sample workloads. The intention of work is to keep the high scaled servers free for longer periods, increasing the chance of handling bulk data process jobs without starving them long. The organization of server farms improves the co-ordination and hierarchical structure of cloud servers.*

Key Words: Cloud Computing, Cloud Server, Load Balancing, Server farms, Load-Score.

1. INTRODUCTION

The emerging technologies in Server Farms improving Cloud Server technology making them robust and efficient in handling client services. The minimization of make span using Ant colony improves more balancing in server farms [1]. The priority based load balancing supported with Genetic Algorithm improved modern cloud computing server farm technology [2] to some extent. Wide range of load balancing algorithms available for cloud computing server farms keeping in consideration of virtual migration of services and load distribution schemes [3]. Two modes of load balancing available for cloud server farms one is based on current state of system called '*Dynamic*' another not based on current state of system called as '*Static*'[8]. The public networks rely on cloud server systems are secured by cryptographic mechanisms such as DES, AES, IDEA and RSA algorithms [4]. Load balancing among server farms must distribute load evenly improving overall performance and reduce response time [5][6]. Application of data structures in scheduling the services for clients adopts efficient features to server farms [7]. In Section 2 Server Farm architecture explained. Section 3 experimental setup described. In Section 4 Cloud based Server Farm computing metrics discussed. In Section 5 methodologies of proposed novel approach is elaborated followed by results, conclusion and future scope of work.

2. SERVER FARM ARCHITECTURE

A server farm or server cluster is a collection of computer servers usually maintained by an organization to supply server functionality far beyond the capability of a single machine. Server farms often consist of thousands of computers which require a large amount of power to run and to keep cool. The typical architecture of Server Farm with a cloud based central server is shown in Figure 1.

2.1 Components of Server Farm

Server Farms can be configured in different ways from business to business. Moreover all Server Farms have similar components with similar setup.

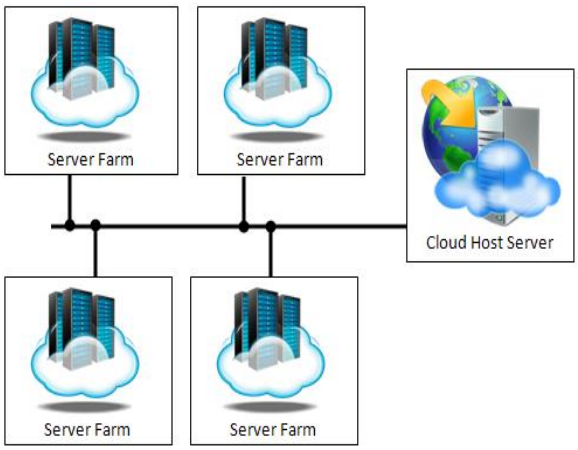

Fig -1: Server Farm based Cloud Host Server

The Server Farm Components are as follows

Fire-Wall

The gateway protector for companies data transfers among Internet applications. Provides support for various security certificates and schemes.

Web Servers

Servers support process of user requests along with SQL database and web service functionalities. High end interfaces are provided with inbuilt secured socket layer implementations to secure business transactions over Internet oriented services.

Databases

Cloud Server Management system organizes databases for Server Farms. High end configurations maintained to support modern database technologies with cryptographic authentications.

Load Balancer

The traditional Load-Balancer balances the load of users across the Server Farms. It directs the traffic of users to company Server Farms. The farms are occupied in a sequential manner and all the farms maintain same configuration and capacity.

3. EXPERIMENTAL SETUP

Software module developed for this experiment using VB.net software to organize three server farms each maintains three co-operative servers. Server farms are configured with capabilities first one normal processing, second one medium and third one high. All the servers hold a central resource governed by host central cloud server. In this experiment file transfer to clients by a cloud server observed. The file size and data transfer rate used as metrics in this experiment. The central resource handled by cloud server organizes a Binary Search Tree data structure to load handling among server farms.

4. CLOUD COMPUTING METRICS

File Transfer Rate

The size of file factor influences the data transfer rate in networks. Smaller the file sizes faster the transfer rate. Relatively large files need more date rate which costs. In corporate sectors and web services huge amount of files interchanged between internet workgroups. Typically file transfer rate represented in Kbps.

$$FTR = \frac{Amount\ of\ File\ transfered}{Number\ of\ files\ to\ transfer}$$

Response Time

The time taken to process client request and allocate of resources to client by a cloud server is response time usually measured in ms (Millie seconds).

Server Scalability

The amount of work load a cloud server /server farm can handle depends on RAMs, CPUs enclosed in server unit, measured in Giga-bits per seconds.

$$ServerScalability = \frac{Total\ Workload}{\sum(RAM\{1..N\}, CPU\{1..N\})}$$

Availability Rate

It is the amount of server time availability to clients for completing their tasks using the services of a cloud server.

$$AvailabilityRate = \frac{Time\ granted}{Total\ waiting\ time}$$

Network Capacity

Network capacity depends upon bandwidth of communication channels and throughput of cloud server. Always cloud servers maintain high capacity networks but clients come in variety of networks with low to medium bandwidths. It is directly proportional to file transfer rate metric.

$$NetworkCapacity = \frac{No.\ of\ bits\ transfered\ per\ sec.}{Total\ no.\ of\ bytes\ to\ transfer}$$

These metrics are used to calculate Load-Score for client request queue for cloud service.

$$LoadScore = \frac{AR}{NC} + \frac{RT}{WT} + (CS_{FTR} - C_{FTR})$$

Where AR=Availability Rate, RT=Response Time, NC= Network Capacity, WT=Waiting Time, CSFTR=Cloud Server FTR and CFTR=Client FTR. Some cloud computing metrics used to generate a score on which load sharing technique depends.

5. METHODOLOGIES

5.1 Proposed Optimal Load Sharing Model

The client file requests are handled by a Balanced Binary search Tree. Hence the client requests are ordered by file size of their request. The server farms are organized into three categories {Normal, Medium and High} based on their processing capabilities as shown in Table 1. Each Server-Farm is a collection of 3 collaborative processors that behave like a single processor and having access to central resource.

Table -1: Client File Transfer Requests

Server Farm	*CPU Capacity*	*RAM*
1	1.2 GHz	6 GB
2	2.4 GHz	12 GB
3	4.8 GHz	20 GB

5.2 Job Selection Methods

The client file requests are ordered based on their file size using a binary search tree. Once the BST generated the upcoming client requests are added as new nodes to BST. The significance benefit of BST is the traversal of tree gives a listing of client jobs in an ascending order of file size. Once the list is ready the following proposed approaches used to process the client requests using server farms.

5.2.1Maxima Minima Selective Method

In this method successive minima from list are always placed into server-farm1 and server-farm2. The maxima from list placed into server-farm3. The client requests which got place in sever are removed from BST and tree re-ordered for each turn. The process always keeps server-farm1, server-farm2 and server-farm3 exchange the process threads and resource handlers according to proposed logic until all the nodes in BST exhausted.

5.2.2 Sequential Selective Method

In this method the nodes in BST are handled sequentially in tree traversal. The client requests are allotted sequentially to server-farms. The technique places the request in front of server-farm immediately when it is idle. The process ends when all nodes of BST are processed.

5.3 Load Score (Ω) Method

The load score evaluated for each job using Eq(1). The load score factor influenced by

$\Omega = \sum_{i=0}^{n} Process\_Time_n + \sum_{i=0}^{n} Data\_Size_n$ ... Eq(1)

Two primary measures one is process time taken by individual tasks in a job, second is data size processed per each task in a job. The jobs are ordered from low to high using the BST as in above method. Using load factor more precisely the server-farms are loaded with jobs. A statistical method can be used to identify more accurate measures to improve the quality of load factor.

6. RESULTS AND ANALYSIS

Table -2: Client File Transfer Requests

Client	*File Size(KB)*
C001	1250
C002	2750
C003	9600
C004	12560
C005	24800
C006	846
C007	1560
C008	21600
C009	13200

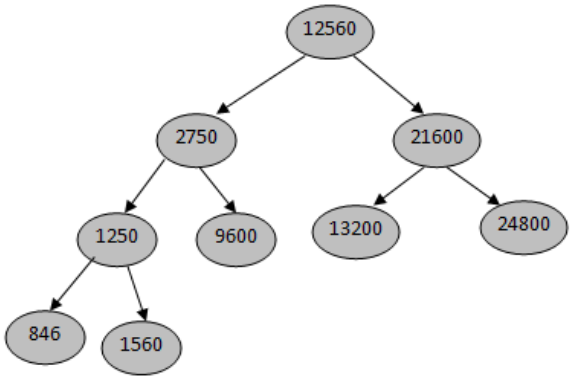

Fig -2: Balanced BST generated on File Size

The software module developed using VB.net to visualize the process of server farm load balancing for the given training sample of jobs. The client requests of file transfers are shown in Table 2. The BST generated depicted in figure 2. The nodes in BST removed and queued into server farms such that the left sub-tree leaf nodes are forwarded to Farm 1&2 and right sub-tree leaf nodes are forwarded to Farm 3. Once a farm becomes free according to its level number BST nodes (client jobs) are loaded into server pools. The BST dynamically balanced each time when a node removed and reduces ambiguity. The jobs from outside still loaded into queue and from it jobs are added to BST as fresh nodes.

Once the BST generated the Process-Monitor module loads jobs into server farms. The module uses load balancing logic and interchanges file transfer loads among three farms. The module simultaneously runs the timer to track the overall process time taken to complete the total jobs in BST. In this experiment a noticeable relief observed on Server-Farm 3 compared to traditional server farm scheme. The distribution of files low to large placed among low to high capacity server farms.

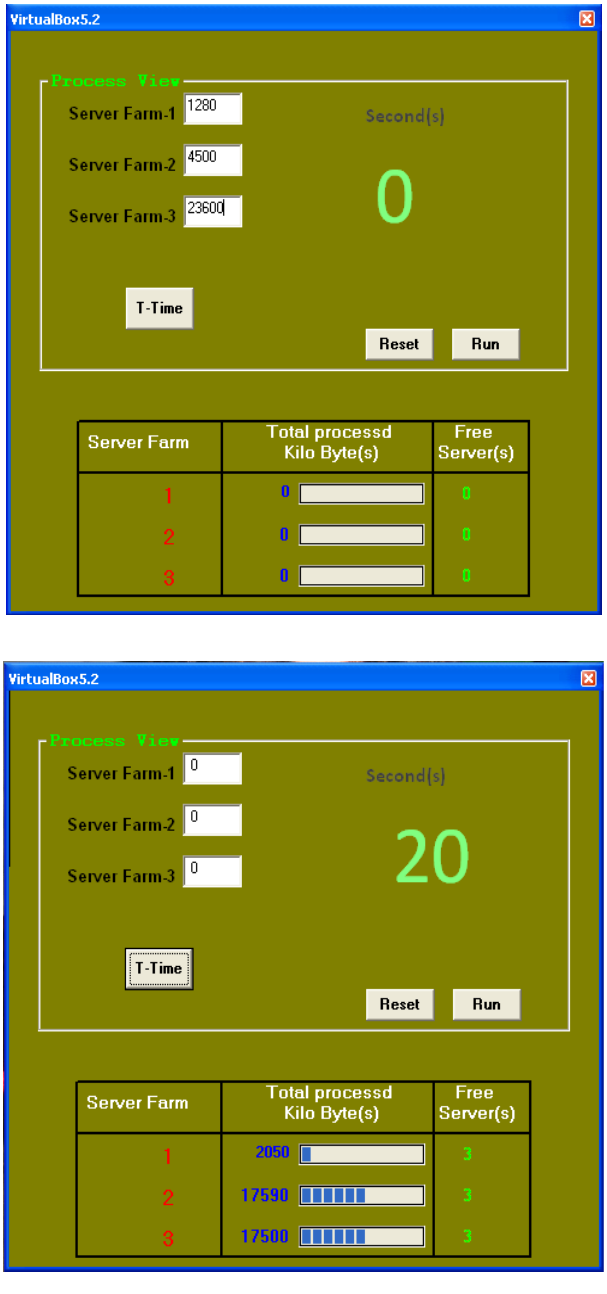

Fig -3: Processor Monitor Snapshot

The experiments conducted by taking three waves of job pools consisting of files to be transferred from server to client. The workload shifting among server farms are closely monitored as shown in figure 3. The Server farm-3 with high configuration is examined over work load distribution. The major intention is to divert short workloads to lower configuration server farms as well as handling high workloads automatically from BST generated nodes using swapping mechanisms.

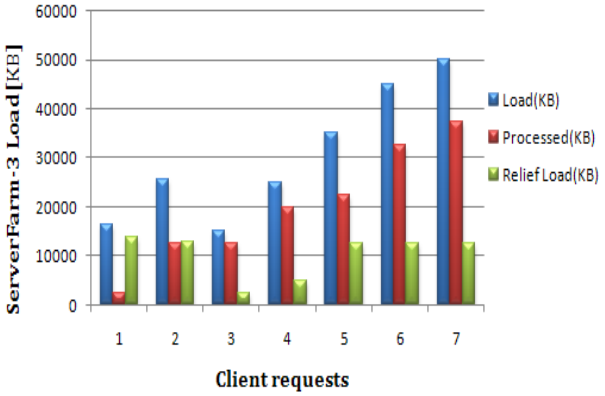

Fig -4: Server Farm-3 Workload Analysis

Table -2: Client File Transfer Requests

Waves	SF-3 released time (traditional)	S F-3 released time (proposed)	Difference
1	145 Sec	56.9 Sec	88.1 Sec
2	226.8 sec	111.2 Sec	115.6 Sec
3	589.4 Sec	348.3 Sec	241.1 Sec

The analysis from the table-2 shows an improvement in Server Farm 3 workload scheduling. The experiments provided support for keeping free Server-Farm3 as soon as possible helps the efficiency of high capacity job handling. The Figure 5 shows the improvement of readiness of ServerFarm-3 among traditional and proposed approaches.

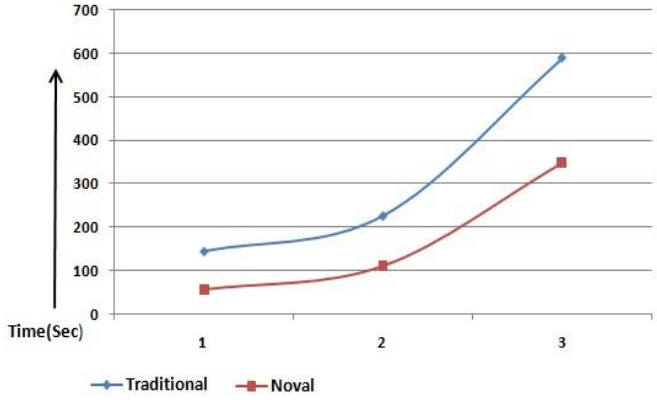

Fig -5: Server Farm-3 Workload Graph

From the above figure-5 an average difference of 148.26 Seconds occurring for Server-farm3 becoming free from workload using novel balanced load sharing mechanism proposed in this paper. It is also easy to implement the BST data structure and logic of inter-swapping among server farms in real-time environments.

5. CONCLUSION

The Novel Load Balancing scheme proposed in this paper reduces the unnecessary waiting of bulk amount file transfer requests to Cloud Server. It keeps track of requests using Binary Search Tree and automates the distribution of requests among Server-Farms. The Inter Swapping of file transfer loads among Server-Farms designed to make High end Server Farms always available to face big jobs. In future the proposed model is to be applied over a real-time network to estimate the model efficiency.

REFERENCES

[1] V.Prachi et. al., "Enhancing Load Balancing in Cloud Computing by Ant Colony Optimization Method", IJCERT, Volume 4, Issue 6, pp. 277-284, June-2017.

[2] Rishav.C et. al., "Cryptography in Cloud Computing: A Basic Approach to Ensure Security in Cloud", IJESC, Volume 7 Issue No.5, May-2017.

[3] Foram F et. al., "Load Balancing in Cloud Computing", IJEDR, Vol.2, Issue-1, ISSN: 2321-9939, 2014.

[4] Md. Shahjahan Kabir et. al., "Process of Load Balancing in Cloud Computing using Genetic Algorithm", ECIJ, Volume-4, Number-2, June-2015.

[5] B.Harmandeep et. al., "A Survey of Load Balancing Algorithms in Cloud Computing", International Journal of Computer Science Trends and Technology (IJCST) – Volume 2, Issue 3, May-June 2014.

[6] J.SamarsinhPrakash et. al., "Load Balancing in Cloud Computing", International Journal of Science and Research (IJSR), ISSN: 2319-7064, Volume-3, Issue-6, June 2014.

[7] Dr. C.S.R.Prashanth et. al., "Review of Load Balancing in Cloud Computing", IJCSI, ISSN: 1694-0784, Vol. 10, Issue 1, No 2, January 2013.

[8] Mayanka.K et. al., "A Comparative Study of Load Balancing Algorithms in Cloud Computing Environment", IJDCC, Volume-1, Issue-2, December 2013.

BIOGRAPHIES

Currently studying B. Tech (CSE) in ISTS-Women's Engineering College. Her areas of Interest are Mobile Computing, Data Mining and Cloud Computing

Research wing member of University College of Engineering, ANRU. He is an author for few publications and his areas of interest are Mobile Computing, Data Mining and Networks.

Currently studying B. Tech (CSE) in ISTS-Women's Engineering College. Her areas of Interest are Mobile Computing and Cloud Computing

Currently studying B. Tech (CSE) in ISTS-Women's Engineering College. Her areas of Interest are Mobile Computing and Cloud Computing

Currently studying B. Tech (CSE) in ISTS-Women's Engineering College. Her areas of Interest are Mobile Computing and Cloud Computing.