ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 6, June 2017

Improved Load Balancing Algorithm for Reducing Idle Iteration in Virtual Machine on Hadoop Distributed File System (HDFS)

G Kanmani[1], E Jayabalan[2]

Research Scholar, Department of Computer Science, Government Arts College, Salem, India[1]

Assistant Professor, Department of Computer Science, Government Arts College, Salem, India[2]

ABSTRACT: The growing technology of cloud computing established for the worldwide users. The every user likes to utilize the delightful cloud services. Because cloud computing provide their services in user friendly as well as some limited amount of cost and its flexibility. In generally cloud functions their mass storage in remote location and accessing for remote server in virtualization technology. To the behavior of cloud at initially developed for store and retrieving for enormous data in virtually. In any type of organization accessing their data at any location based on the internet. Cloud computing mass storage server maintenance is critical task since the developing users are increase for every day by day. Load balancing is vital role for the cloud computing. Without load balancing is critical task for accessing cloud services. In much organization provide number of features based rendering services. The proposed algorithm is support for proper uploading files and scheduling the files in the server simultaneously. To the scheduling is based on huge data storage via the load balancing level and finally distributing load to the consumers as well as customers. This approach stands for mainly reducing task idle iteration in the database and minimizing response time for comparing physical machine into virtual machine.

KEYWORDS: Load Balancing; Virtualization; Online Mode Heuristic Algorithm(OMHA); Map reduce,Skewness, Clustering, Scheduling, Idle Iteration.

I. INTRODUCTION

The user friendly technology of the cloud computing is formulated by the different kind of services. The individual cloud computing services offering organization is covered for some area belongs to the security restrictions for its storage. The high availability occurring to the server should definitely need for cloud load balancing. Cloud Load balancing is help for resource availability increasing in to the Physical Machine (PM) and to avoid single machine overload for enhancing system performance [1].

In [2], the digital industries it may be a government or private sectors are using cloud technology. The modern trend in business organizations retrieves the cloud services as well as large mediums like as social media. The internet based cloud technology is derived from a parallel processing and distributed computing. Mainly cloud technology followed by the virtualization technique for effective management in cloud services. Cloud faced many critical aspects issue but every aspect are controlled for balancing the load because majorly constrained load balancing is top most research area in cloud.

Virtualization is essential technique of cloud computing. Physical machine is converting and split into a multiple virtual machine for assigning a multiple task for virtualization technology. The software framework of a physical machine is called as virtual machine [2,3]. Virtualization is divided into two sub parts. One is Container based virtualization which is based on operating system level split in to the memory. Another one is Hypervisor based virtualization which is based on undertakes at hardware level [3].

Load balancing is described at the two levels. The first level considers at Static and second level is Dynamic. Resource allocation and task scheduling is key challenge for effective load balancing. Data handling to the cloud providers should need for load balancing. The amount of work dividing into a physical machine to perform among multiple virtual machines and cluster finally collect the work faster via the load balancing [4].

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 6, June 2017

The cloud load balancing service offering in real time organizations are like IBM, Alibaba Cloud, Amazon EC2, Google. That the organization gives user for rendering the services with monthly cost and also trial day's service also offering for recognizing the activities for store files and load management based on the balancer.

In [4] parallel processing the huge data to be accomplished by the single task and split into multiple independent tasks for data handling. Map reduce is one of the programming model for activating parallel processing for handling large dimensions of data. The Map reduce perform at a cluster level structure and unstructured data to the implement of map, shuffle and reduce stage respectively in parallel.

II. LITERATURE SURVEY

AbhinavThorat, et.al [5], proposed Energy Efficiency Algorithm providing 1) to check the lightly loaded nodes, 2)heavily loaded nodes and 3) idle nodes in the network 4) reassigning the heavy node to light node and idle node for performing their task. This article describing major concern about resource utilization with depends on reliable and securable.The proposed algorithm shows three stage processes. First one task or application is migrating on virtual machine in to server. Second one is finding idle or sleep server for assign task. Final one is reassigning task are migrating into the server on demand. The migrating work can be implemented for increasing throughput and reducing the reply time.In[6], suggested to effective load balancing is depending on redistribution. The heavy loaded node should need to distribute for light loaded node because only managing is possible for the cloud. Load balancing algorithm divide into Batch Mode Heuristic Algorithm (BMHA) and Online Mode Heuristic Algorithms (OMHA) respectively. The suitable cloud environment for needs better results on using OMHA. The algorithm describes honey bee method embrace the directed acyclic graph technique for load transfer. Finally IBEE method reduces the work load in effective for 10-12% comparing between BEE.

The proposed work, requirements are utilizing resources with depend on virtualization technique. The skewness is merging several tasks into server. The cloud using resources fully based on cloud users and cloud providers for their utilization based cost. Overcomingfor the new approach for finding the serverif busy state or not and to assign a task in virtual machine. When hotspot exceed the work certainly intimate the server for heavily loaded because change to virtual machine need to another one. In the method of green computing using for reduce processing power consumption and increasing throughput results with and without load balancing algorithm [7].

In [8], Cloud computing approach is analyzed by the authors, consider at metrics for migration, energy consumption and resource utilization. The paper invented for load balancing and diminishes leakage of resources. The proposed work crossover the three stages. One is selecting a physical machine. Second one is selecting a virtual machine. Final one is selecting physical machine for placing a virtual machine. The results are analyzed CPU, RAM, and bandwidth for using CloudSim Simulator.

In []9], the load balancing is depend on price. Because the scheduling the load approach should base on the priority. In generally most paid users are gives high priority. The cost conscious scheduling is followed by balancers. The proposed genetic algorithm faced on three stages namely Selection, Operation and Replacement for justifying the load. The genetic algorithm via the load balancing performs at location rule, distributing rule and lastly selection rule. The results discussed for response time for data centers in effective. Theauthor [10], consider about load balancing goals are Cost, scalable, flexible and priority. The experimental results show that response time, processing cost and process request service time. Here taken that 3 data centers, 6UserBases and finally 25, 50, 75 virtual machine respectively for the implementation.

IV. METHODOLOGY

This proposed hybrid approach is based on skewness algorithm and map reduce programming model. That is one kind of Online Mode Heuristic Algorithm (OMHA). The hybrid approach is dynamical load balancing technique for cloud computing. The IBM offering number of services for accessing cloud services to deploying the application. The few IBM services are cost and many are free. The work developed by the trial IBM Bluemixservices with support of 5GB data storage. Initially the work evaluated physical machine and balancing the load in virtual machine on utilizing IBM services. The BigInsight apache hadoop, Object Storage and Availability Monitoring Services are used for effective load balancing. The flow of work is based on organizational representation in cloud.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

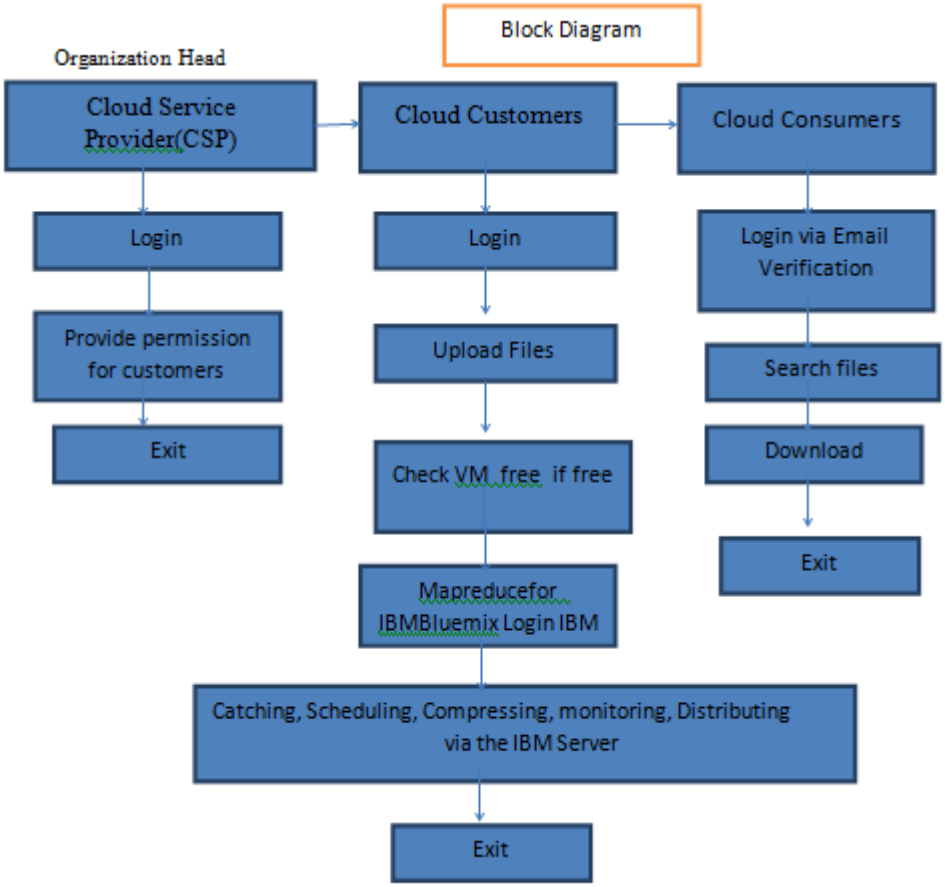
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 6, June 2017

Description: The hybrid approach is fully based on organizational research in cloud environment. How cloud computing works in IT environment is shown in this approach. The cloud platform is very critical task to load management that situation is overcome by proposed method. Each cloud regions are allocated different migrated servers for storing data. The same data can split into multiple places for security concerns. If it user needs that they all are collect from a different location in the cloud that situation handling for scheduling the files in order to clustering. The cloud service provider given that permission only customers access their cloud services. The cloud Service Provider have monitoring all transferring data if it is upload or download by the customer as well as consumer. The customer uploading files are checked to cloud servers if it idle or overload. The overload server is called hotspot and idle server is considering at coldspot. So the files are migrating overload server into Idle server using skewness(catches) algorithm. To storing files are recollecting process are used by mapreduce programming method via the hadoop distributed file system for proper balancing(compress). The proposed algorithm cloud service for accessing in Software as a Service via the IBM Bluemix for cloud customers. The parallel processing of storing and balancing is provide by cloud consumers. The load management and distributing all are given effective performance in this hybrid algorithm.



The load balancing in the large files handling in hadoop framework is reducing the idle iteration, increasing a threshold level and response time are calculated in the work. The uploaded files are monitoring various cloud regions. Then the files are search in the user request and easily download their documents.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 6, June 2017

V. EXPERIMENTAL RESULTS

In generally load balancing metrics are throughput, response time, migration time, fault tolerance are calculated for effective load management. The proposed approach additionally include for reducing iteration and increasing threshold level for suitable load balancing. The uploading cluster files are stored in different remote location to the load balancing in virtually.

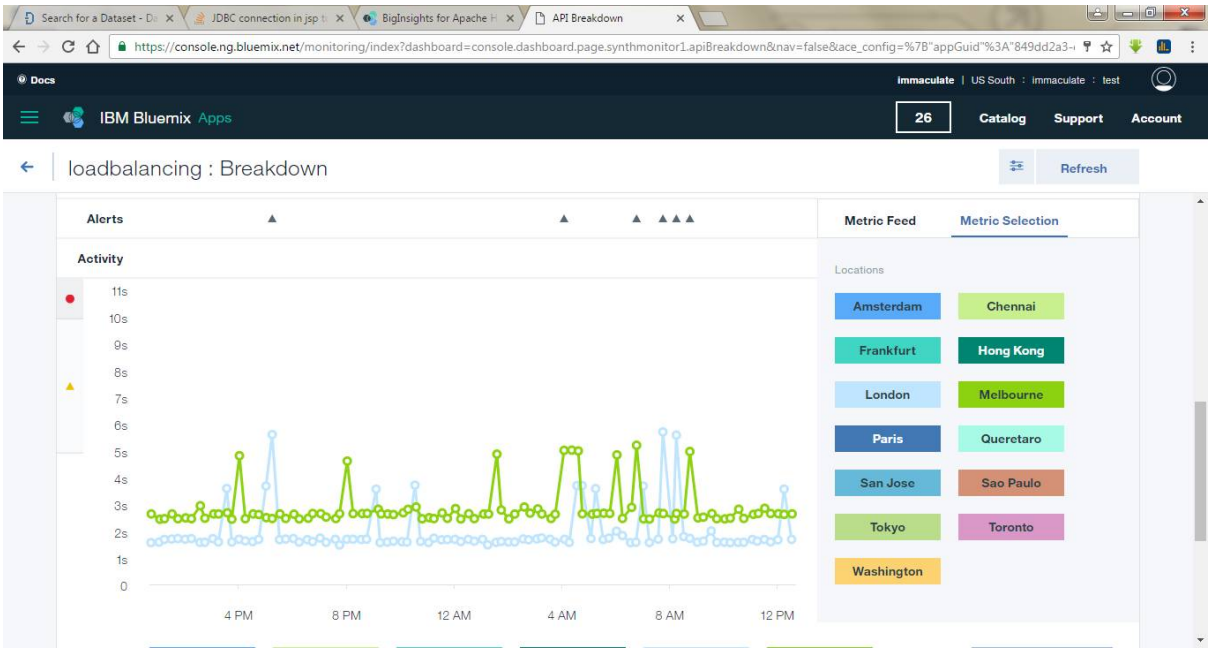


Fig.1. The files stored in virtual machine in different regions

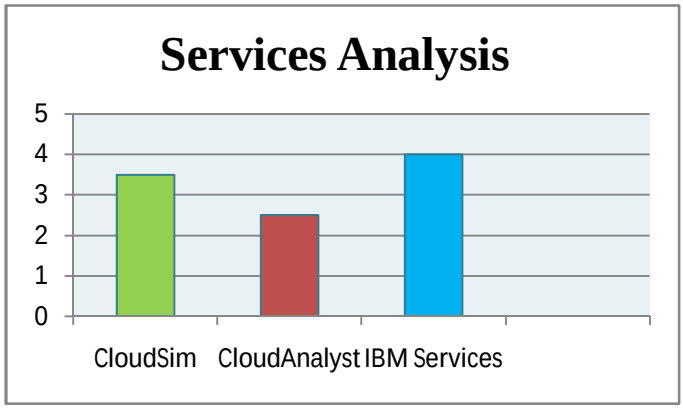


Fig.2. IBM services Best for Comparing Cloud Simulation tools

The most of the researchers to develop a cloud environment is used for cloud simulation tools like CloudSim and Cloud Analyst. In comparing the cloud simulation tool the IBM services is provide better results in fig.1 and fig.2. The load balancing is also done via the IBM services with effectively.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 6, June 2017

Metrics	Physical Machine (PM)	Virtual Machine (VM)
Idle Iterations	high	low
Response time	high	low
Throughput	low	high
Migration time	low	high

Table 1. Load Balancing Metrics Compare between PM to VM

VI. CONCLUSION

The cloud load balancing is mainly based on resource allocation and scheduling task. With the IBM services using their flexible services for effective load management. In the proposed approach the files are stored in parallel and balancing the load in simultaneously. The response time increased for comparing physical machine into virtual machine and idle iterations are reduced. Scheduling the load is properly managed via the map reduce programming model. In the future enhancement the cloud load balancing fully utilize cloud local and global load balancing with appropriate cloud services.

REFERENCES

1. Cloud Load Balancing-Wikipedia.
2. Navtej Singh Ghumman, Rajesh Sachdeva, " An Efficient Approach for Load balancing in Computing using Composite Techniques" , International Journal of Research in Engineering and Applied Sciences, Volume 6, Issue 2 (February, 2016).
3. Mamtakhanchi, Sanjay Tyagi, " An Efficeient algorithm for Load Balancing in Cloud Computing", International Journal of Engineering Sciences& Research Technology, June 2016.
4. Vibha M B, Dr. R.RajuGondkar, "Smart Load Balancing towards Effective Cloud Computing", International Journal of Advanced Research in Computer and Communication Engineering, Volume 5, Issue 6, April 2016.
5. AbinavThorat, ShrinivasSonkar, "Energy Aware Load Balancing Technique For Managing Network Workload in Cloud Computing", International Journal of Computer Science Trends and Technology, volume 4, issue 3, May-June 2016.
6. Abhishek Kumar Tiwari, Sreeja Nair, " Improve the Efficiency of Load Balancing in Cloud Environment using DAG and Honey Bee Algorithm", International Journal of Computer Science and information Technologies, Volume7(2), 2016.
7. Radha, Dr.Vivek Kumar, " Efficient VM Load Balancing Algorithm for Dynamic Allocation of Resource in Cloud Computing Environment" , International Journal of Innovative Research in Computer and Communication Engineering, Volume 4, Issue 6, June 2016.
8. SachinSoni, Praveen Yadav, "A Load Balancing Approach to Minimize the Resource Wastage in Cloud Computing", International Advanced Research Journal in Science,Enineering and Technology, Volume 3, Issue 3, March 2016.
9. Md. ShahjahanKabir, Kh. MohaimenulKabir, Dr. Rabiul Islam, "Process of Load Balancing Using Genetic Algorithm", Electrical & Computer Engineering: An International Journal(ECJI), Volume 4, Number 2, June 2015.
10. Dr.S. Suguna, R. Barani, "Simulation of Dynamic Load Balancing Algorithms", Bonfring International Journal of Software Engineering and Soft Computing, Volume 5, No.1, July 2015.