Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 1189 - 1195
Received 20 January 2021; Accepted 08 February 2021.

Review and Analysis Load Balancing Machine Learning Approach for Cloud Computing Environment

Authors Name: Afreen Bari
Department: Computer Science & Engineering
Shri JagdishprasdJhabarmalTibrewala University
Rajasthan, India
afreenabdulbari@gmail.compeddiprasad37@gmail.com

Guide Name: Dr.Prasadu Peddi
Department: Computer Science Engineering
Shri JagdishprasdJhabarmalTibrewala University
Rajasthan, India

ABSTRACT

The discussion centres on load balancing approaches utilising the Artificial Learning Approach cloud computing paradigm. Cloud load balancing is a series of activities to distribute workloads to underused VMs, such that resources can be more efficiently distributed across cloud computing. In this paper, we discuss a Approach to design a load handling system with a similar case study to incorporate various scheduling approaches. The design of the model will be carried out using a managed approach that chooses to test log data on the basis of predictions. It maximises resource utilisation in different condominiums, such as low, medium and peak loads, and also provides device reliability. This induces a dynamic system environment.

I. INTRODUCTION

To maintain device stability, reduce critical response time and increase resource performance, use load balancing strategies to dynamically delegate demands that enter the cloud setting. One of the main challenges for complex load handling is an improvement in interim communication overheads (swapping files between VMs). The problem of overhead connectivity is neglected in many of the methods proposed for load balancing. The goal of the autonomous load balancing device is to solve this issue. CPU-bound demands are also based on the accessible Cloud System Task Planning studies. In this case, the parameters are broken down into Processor and I/O-bound requirements based on the resources. Taking into account all types of specifications, the open type of load balancing cannot be enforced. This suggested algorithm would distribute the working burden between them equally and assign applications to the applicable VMs on the basis of the required resources; thereby minimising overhead and increasing the degree of load balancing. This article proposed a hybrid optimization algorithm and a bee colony algorithm. This analysis touches on the key challenges and issues of cloud computing. Cloud computing is the most state-of-the-art information infrastructure still providing customer support. LB is one of the key worries of the CC as a device overload. An appropriate LB algorithm for the efficient use of resources is therefore also needed. The main priority of LB is the fulfilment of customer needs by the distribution of workload through multiple network nodes and the maximisation of resource usage and improved device efficiency. Efficient load management is also critical to the efficiency of the service.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 1189 - 1195
Received 20 January 2021; Accepted 08 February 2021.

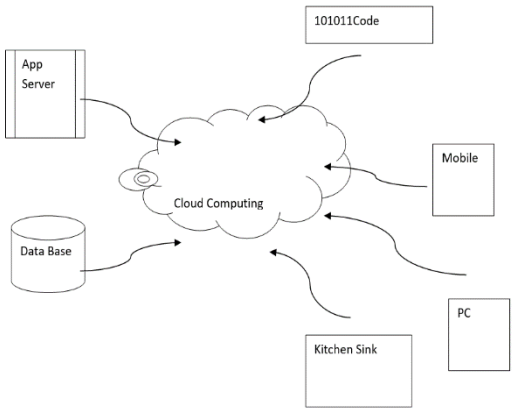

Figure 1: Basic cloud computing approach

This study deals with the numerous LB load balancing algorithms, the hierarchical load balancing algorithm, and the form load balancing dynamics. The necessity to create fully autonomous new, sophisticated LB algorithms in the future will allow for better usage of energy, reduced space and enhanced levels of inadequacy, efficient migration of tasks and a short amount of time. The CC itself is a long-term technology. It is one of the most relevant innovations that we can use to deliver on a core pillar of our market. In the long run, the above advances will give CC maximum impetus. The mechanism suggested shall take account of the characteristics of a complex network. The methodology is coupled with the establishment of the new ACO method for the refurbished ant colony algorithm.

II. RELATED WORK

Son and E. Huh [1] This paper presents the QoS metrics of the migration system and proposes an unnecessary migration justification scheme. The main emphasis of this paper is the application of a fleeting rationale and a sophisticated migration machine learning approach.

S. Paul et al[2] In this paper, we have devised a dynamically balanced loading strategy based on the simple neural network technology that will dynamically rank servers based on the rest of the server load capacity and deploy the assignment to the optimally loaded server in the best-fit virtual machine instances. The machine learning system for load power.

Q. Liu, et al[3] MapReduce is a standard paradigm for distributed programming. Load balancing is key to the effect of distributed computing performance in large-scale systems, such as home energy storage in villages, online virtual communities, etc. The suggested MapReduce load balancing approaches are structured to optimise job runtime, but disc space is not taken into account. This paper introduces a new arrangement made up of updated K-ELM and NSGA-II. Related experimental results have shown that we can assign roles fairly and maximise the performance of a cloud device.

A. Mathew et al[4] The aim of the present paper is to simulate the current, perform time series analyses using neural networks using pre-collected data sets and forecast the necessary performance, as well as an intelligent optimization of the data spectrum.

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 1189 - 1195
Received 20 January 2021; Accepted 08 February 2021.

J. Zhong et al[5] Our experiments were carried out in the Amazon EC2 8 node virtual cluster. Our preliminary results indicate that 1) the GPU is a viable cloud-based graph processing accelerator and 2) the suggested optimizations will improve the performance of the GPU graph processing engine. The experiences we have obtained and the problems of large-scale GPU acceleration graphical processing are addressed.

M. Sultan et al[6] Propose the application of the ML approach to collect statistics from networks and to establish an intelligence policy. We show that this data centre optimization plane is feasible with SDN controllers and Southbound APIs. In order to model future rises and declines, we define the architectural components of this technique – using a network simulator. We already have open source software that can be used to provide the scalable capacity of a data centre based on load predictions.

R. Khanna, et al[7] We can explain the system by learning and classifying workload runtime actions to create a multi-phase model. We use a variant of the machine learning approach to demonstrate the concept of step-by-step workload estimation. The findings indicate that the latest model in phase detection is approximately 98 per cent reliable and 97.15 per cent accurate in estimation of demand forecast.

III. LAD BALANCING IN CLOUD ENVIRONMENT

Cloud computing is inexpensive and flexible, but retaining processing stability ensures that many cloud computing workers will contribute to an awful loss.

load balancing allows the optimal usage of available resources and at the same time minimising power consumption.

The knowledge is transmitted and retrieved without any noticeable delay by splitting the traffic between the servers. Various types of algorithms that support load traffic between open servers are available.A instance of load balancing in our lifestyle is related with websites. Without load balancing, users might expertise delays, timeouts and doable long system responses.

Depending on how the charge is distributed and how processes are allotted to nodes i.e., System Load.

- Depending on the data status of the nodes i.e., System Topology.

In the first instance, this method is designed as clustered, dispersed or hybrid and in the latter as static, dynamic or reconciliative.

The key goal of LB is to handle load effectively through multiple cloud nodes such that nodes are not under/overloaded[7]. In order to maximise the usage of infrastructure to minimise overall response time, LB may be defined as a method of distributing burdens through network connections on several devices or systems clusters. It decreases the overall waiting time of the system and thus prevents unnecessary reproduction of properties. During this method, requests disperse inside servers to transmit and process data without waiting. LB is the approach by moved computer load to optimise machine efficiency.[9]

LB proposes a systemic framework for fair allocation of liability for sufficient capital. In the event of a catastrophe in part of the service, the purpose is to provide reliable service, including sufficient usage of the

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 1189 - 1195
Received 20 January 2021; Accepted 08 February 2021.

resource, by providing and disassociating the equipment case. In addition, LB aims at reducing the reaction time for activities and increasing the utilisation of services. [9].

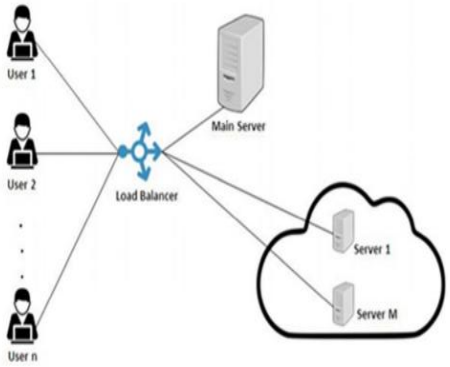

Figure 2: load balancing in cloud environment

IV. TYPES OF LOAD BALANCING ALGORITHM

a) STATIC LOAD BALANCING

This involves a preliminary awareness of the systems capabilities in order not to rely on the current state of the system for the determination of shifting the load. The machine operates with a minimal load variance.

b) DYNAMIC LOAD BALANCING

Here the current state of the system is used to make decisions to manage the load.

- Distributed
- Non-Distributed.

DISTRIBUTED DYNAMIC LOAD BALANCING

The dynamic load balance algorithm is performed in the distributed one by the entire nodes of the scheme which share the load balance role. The contact between load balancing nodes will create two forms: cooperative and non-cooperative.

Dynamic load balancing of disseminated character algorithms also produces more messages than those not transmitted, since each node of the device has to communicate with all other nodes. An advantage of this is to avoid the complete load balancing operation, even though a single or more nodes in the device fail. Instead, it would especially upset the scheme's efficiency.

Distributed dynamic load balancing will familiarise a framework that allows any node to share status information with each other's node.

NON-DISTRIBUTED DYNAMIC LOAD BALANCING

In non-distributed type, whichever unique node or a collection of nodes do the job of load balancing. It canister two forms:-

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 1189 - 1195
Received 20 January 2021; Accepted 08 February 2021.

- Centralized: a single node in the whole device performs the load balancing algorithm: the key Node. This node is liable for the entire system's load balance
- Where the load balance is centralised in each cluster. In each cluster, a central node is selected by an acceptable form of election to ensure the load balance within this cluster.

CENTRALIZED LOAD BALANCING

In centralized load balancing [3] The method is complemented by a single node in both the provision and scheduling judgement. This node is responsible for storing the entire cloud network's data base and can connect static or dynamic load balancing approaches. This approach decreases the amount of time it takes to evaluate multiple cloud services but provides the unified node a summary.

In this situation the network is often not an expanded fault resistant, since the overwhelmed clustered node's loss concentration is strong and the recuperation power in the event of node failure is not casual.

HIERARCHICAL LOAD BALANCING

It includes the separate phases of the cloud in the decision to match loads. This load balancing techniques mostly operate in master slave mode. The expenditure tree data structure indicates that each node in the tree is balanced below its parent node supervision.

MACHINE LEARNING FOR LOAD BALANCING IN CLOUD COMPUTING

This paper presents numerous scheduling algorithms, such as the Cat Swarm Optimization Algorithm and Cockoo Quest. This paper presents various ML scheduling algorithms such as the Virtual Annealing Algorithm.

1) Simulated Annealing: Simulated annealing is influenced by the rinsing of solids, where the flushing of solids, such as metal or glass, requires heating, which gradually cools and reduces internal stresses. Usually, this method, locked with local constraints, consists of needless, qualified assignments, and the algorithm always depends on the nature of the request and the volume of the bin. The system performs well at high temperatures.

2) Tabu Search: The Tabu searchalgorithm is a global optimization method that has been built to emulate human intelligence and is more capable of maximising consistency. It tries to direct other approaches to break the ideal spatial trap and is used to fix the allocation of resources and to optimise other problems.

3) Partical Swarm Optimization: Extremely sophisticated heuristic algorithms for smart optimization, copying animal swarming behaviour. The PSO algorithm is not efficient in solving differential problems. PSO successfully overcomes a variety of issues related to pair optimization due to its concurrent allocation, extension, quick-to-understand and efficient longevity with outstanding functionality in complex environments.

V. Conclusion and Future work

cloud computing is a dynamic load balance, with a large spectrum of studies and one of the main study topics being cloud computing. On request for service, the customer should be available to the user. If the node is overloaded at this stage, the load balancer must set the load to a different free node. Load balancing in cloud computing is often very critical. Therefore, the following review will concentrate on algorithms, first, loading on the server and then on the actual effects of the server.. We will implement in near future The ACO algorithm

Annals of R.S.C.B., ISSN:1583-6258, Vol. 25, Issue 2, 2021, Pages. 1189 - 1195
Received 20 January 2021; Accepted 08 February 2021.

can be used in clouds to enhance efficiency, resource use and energy usage, but more can also be used to manage cloud loads, and performance can also be increased with different parameters that in this research project.

REFERENCES

[1]. Son and E. Huh, "Study on a migration scheme by fuzzy-logic-based learning and decision approach for QoS in cloud computing," 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN), Milan, 2017, pp. 507-512, doi: 10.1109/ICUFN.2017.7993836.

[2]. S. Paul and M. Adhikari, "Dynamic Load Balancing Strategy based on Resource Classification Technique in IaaS Cloud," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, 2018, pp. 2059-2065, doi: 10.1109/ICACCI.2018.8554440.

[3]. Q. Liu, W. Cai, J. Shen, D. Jin and N. Linge, "A load-balancing approach based on modified K-ELM and NSGA-II in a heterogeneous cloud environment," 2016 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, 2016, pp. 411-412, doi: 10.1109/ICCE.2016.7430670.

[4]. A. Mathew, N. E. I Deepu and M. Mohan, "Intelligent Edge Security with Dynamic Task Offloading in Fog Environment," 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2019, pp. 367-372, doi: 10.1109/ICCES45898.2019.9002417.

[5]. J. Zhong and B. He, "Towards GPU-Accelerated Large-Scale Graph Processing in the Cloud," 2013 IEEE 5th International Conference on Cloud Computing Technology and Science, Bristol, 2013, pp. 9-16, doi: 10.1109/CloudCom.2013.8.

[6]. M. Sultan, D. Imbuido, K. Patel, J. MacDonald and K. Ratnam, "Designing knowledge plane to optimize leaf and spine data center," 2020 IEEE 13th International Conference on Cloud Computing (CLOUD), Beijing, 2020, pp. 13-15, doi: 10.1109/CLOUD49709.2020.00011.

[7]. R. Khanna, M. Ganguli, A. Narayan, A. R. and P. Gupta, "Autonomic Characterization of Workloads Using Workload Fingerprinting," 2014 IEEE International Conference on Cloud Computing in Emerging Markets (CCEM), Bangalore, 2014, pp. 1-8, doi: 10.1109/CCEM.2014.7015482.

[8]. Patel, D., & Rajawat, A.S. (2015). Efficient Throttled Load Balancing Algorithm in Cloud Environment. International Journal of Modern Trends in Engineering and Research, 2.

[9]. M. Ibrar et al., "ARTNet: AI-based Resource Allocation and Task Offloading in a Reconfigurable Internet of Vehicular Networks," in IEEE Transactions on Network Science and Engineering, doi: 10.1109/TNSE.2020.3047454.

[10]. Li, M., Zhang, J., Wan, J. et al. Distributed machine learning load balancing strategy in cloud computing services. Wireless Netw 26, 5517–5533 (2020). https://doi.org/10.1007/s11276-019-02042-2

[11]. Bui K.T., Nguyen L.V., Tran T.V., Pham TV., Tran H.C. (2021) A Load Balancing VMs Migration Approach for Multi-tier Application in Cloud Computing Based on Fuzzy Set and Q-Learning Algorithm. In: Kumar R., Quang N.H., Kumar Solanki V., Cardona M., Pattnaik P.K. (eds) Research in Intelligent and Computing in Engineering. Advances in Intelligent Systems and Computing, vol 1254. Springer, Singapore. https://doi.org/10.1007/978-981-15-7527-3_58

[12]. Vatsalya S.P., Vidhya Sree N., Chethan Malode C.M., Bhargavi K. (2020) Apprenticeship Learning Based Load Balancing Technique for Cloud Environment. In: Pandian A., Ntalianis K., Palanisamy R. (eds) Intelligent Computing, Information and Control Systems. ICICCS 2019. Advances in Intelligent Systems and Computing, vol 1039. Springer, Cham. https://doi.org/10.1007/978-3-030-30465-2_74

[13]. Ghasemi, A., Toroghi Haghighat, A. A multi-objective load balancing algorithm for virtual machine placement in cloud data centers based on machine learning. Computing 102, 2049–2072 (2020). https://doi.org/10.1007/s00607-020-00813-w.

[14]. Polepally V., Shahu Chtrapati K. (2018) Energy Aware GSA-Based Load Balancing Method in Cloud Computing Environment. In: Reddy Edla D., Lingras P., Venkatanareshbabu K. (eds) Advances in Machine Learning and Data Science. Advances in Intelligent Systems and Computing, vol 705. Springer, Singapore. https://doi.org/10.1007/978-981-10-8569-7_29

[15]. Lakshmi S.N., Paul M., Nagaprasad S. (2021) A Nature-Inspired Solution to Managing Activities in the Cloud with Equal Time by Using Machine Learning Approach. In: Chaki N., Pejas J., Devarakonda N., Rao Kovvur R.M. (eds) Proceedings of International Conference on Computational Intelligence and Data Engineering. Lecture Notes on Data Engineering and Communications Technologies, vol 56. Springer, Singapore. https://doi.org/10.1007/978-981-15-8767-2_20

[16]. Tiwari P.K., Joshi S. (2016) A Review on Load Balancing of Virtual Machine Resources in Cloud Computing. In: Satapathy S., Das S. (eds) Proceedings of First International Conference on Information and Communication Technology for Intelligent Systems: Volume 2. Smart Innovation, Systems and Technologies, vol 51. Springer, Cham. https://doi.org/10.1007/978-3-319-30927-9_36

[17]. A. Singh Rajawat and S. Jain, "Fusion Deep Learning Based on Back Propagation Neural Network for Personalization," 2nd International Conference on Data, Engineering and Applications (IDEA), Bhopal, India, 2020, pp. 1-7, doi: 10.1109/IDEA49133.2020.9170693.