International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-9 Issue-2, July 2020

Dynamic Load Balancer for Traffic Management in Cloud Environment

Prithi S, Roxanna Samuel, Ponmani S, J.Jinu Sophia

Abstract: ***The Cloud traffic is the major issue faced by users' every day. Users are not able to access the files at the desired time due to the delay, caused by traffic. Traffic is induced when many users access the same network at a given point of time. This paper aims to reduce the cloud traffic by allowing the user to enter the datacenter for quick access of files with maximum capacity left instead of entering the datacenter in sequential order as in existing systems. The foremost key and challenging problem for handling big data centers in clouds are to balance the Load while flow scheduling since a huge amount of data are transferred at regular intervals throughout a thousand of customers and clients. With a rapid growth in applications, capability of utilizing the data centers has become a challenging task to cloud service, particularly during the peak time usage of data centers and when the requests of user's are unbalanced and when there are amount of demands need to be handled. In this project, Software Defined Network (SDN) controller is applied to improvise the utilization of bandwidth of Dynamic Circuit Network (DCN) as well as reduce the delay that occurs for end users. This project presents the Genetic load balancing algorithm to provide the provider with a high utilization of bandwidth and to low the end-users dealy.***

Keywords: ***Cloud traffic, Load balancer, Datacenter, Load balancing algorithm.***

I. INTRODUCTION

In video streaming such as Netflix, social network such as Face book and web search such as Google which are applications of internet in large scale, it provides more than hundred million services to end users. Because of the massive raising in the application demand, there is a motivation for datacenter that is distributed geographically with enhanced performance and reliability. This resulting applications' deployment lead to problem of request allocation and so the environment requests of end users must be sent to the specific data center. [1]

Manuscript received on May 25, 2020.
Revised Manuscript received on June 29, 2020.
Manuscript published on July 30, 2020.
* Correspondence Author
Prithi S, Assistant Professor, Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India.
Roxanna Samuel, Assistant Professor, Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India.
Ponmani S, Assistant Professor, Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India.
J.Jinu Sophia, Assistant Professor, Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, India.

As an initiative, a straightforward approach of request allocation could be accomplished by allocating each of the requests to its datacenter which is closest.

This native method leads to major two consequences from end user and provider perspectives. First, because of over subscription, the data center will be overloaded during the time of peak workload. Moreover, the datacenters that are different will reach the peak workload in different times. [3,8]

As a result, some datacenters become overloaded and some other data centers may experience bandwidth that is extremely low at a particular time.

For datacenter with low utilization bandwidth, it is waste of energy and investment for provider. [10] Meanwhile this overloading can lead to failures and further to poor performance of application. In order to model utilization in high bandwidth which is providers requirements at all end users and data enters that have low delay, a load based genetic method is proposed. Usually the delay has the transport delay which is outside data center and response time which is inside data center. Then the problem of optimization is formulated. In order to solve it an algorithm called request allocation which will be efficient is proposed by adding the genetic techniques' advantages. The uniqueness and existence of this solution and convergence of the algorithm is theoretically proven. Simulation experiments in workload traces of real world demonstrate that algorithm of request allocation can improve efficiently the utilization of bandwidth for provider, reducing the end users delay and it out performs the locality and greedy algorithms. [6,9]

II. LITERATURE SURVEY

UNREELING NETFLIX was proposed to comprehend and improve multi-CDN movie delivery [2]. Netflix is the chief supplier of subscription-based video streaming service within the North American countries, out of which United States of America accounts to 29.7% of the highest downstream traffic. The indulgent of the Netflix design and its performance will aid in optimizing its style and on the plan of comparable subscription-based video streaming services. This paper makes a measurement study on Netflix to locate its structural design and service strategy [7]. It is found that Netflix employs combination of datacenters and Content Delivery Networks (CDNs) for content deliverance. It also performs live measurement of the three CDNs used by Netflix to determine the video delivery bandwidth accessible on the market to users across the North America. Finally, as an enhancement to existing CDN assignment strategy in Netflix, it was proposed to include a measurement-based adaptive CDN selection strategy and a multiple-CDN-based video delivery strategy.

This addition confirmed a considerable increase in user's average bandwidth.

An open access, scalable, shared geocast service for distributed systems was proposed by author in [4] called CLOSESTNODE.COM. ClosestNode.com is a specific, expandable and backwards-compatible service for outlining clients to the closest server. It supports a DNS interface that helps the clients to find a service name, and gets the IP address of the nearest server.

A group of shared systems are used for mapping which minimizes the administrative and execution expenditure of proximity-based server preference.

It is also designed toward reducing the cost of effort needed for system developers to form novel framework service which at the nearest vicinity.

The collaboration of the network in multi-tenant datacenters will considerably influence the network performance of the tenants, which will change drastically. This will relate not only to production datacenters but also cloud environments. Network performance inconsistency will also influence the application performance which in turn causes erratic tenant cost and loss of suppliers' income. Motivated by these factors, the authors in [5] tried to broaden the tenant-provider interface to more evident account for the network. This can be achieved by providing tenants with a virtual network which connects to their instances. The primary aim is to design a virtual network abstraction which guarantees performance to the tenants, their worth and the provider returns. To demonstrate the possibility of virtual networks, an Oktopus system was developed to implement the proposed abstractions. By means of extensive simulations and an Oktopus operation on a 25-node two-tier test-bed showed the usage of virtual networks which was proves to yield enhanced and predictable tenant performance. Moreover, employing a straightforward valuation model reduces tenant rate by up to 74% while nurturing the providers' income fairly.

III. ARCHITECTURE DIAGRAM

Specifically, primarily based methodology was used to design a load balancer that meets the providers requirement of increasing the information measure usage in the least datacenters and to reduce end-users delay. The optimization problem is considered as the request allocation formulated for the providers requirement. To solve it, an associate degree economical request allocation algorithmic rule was proposed by introducing the genetic technique to remove difference, in contrast with directly implementing the nash algorithm approach which is depicted in Figure 1.

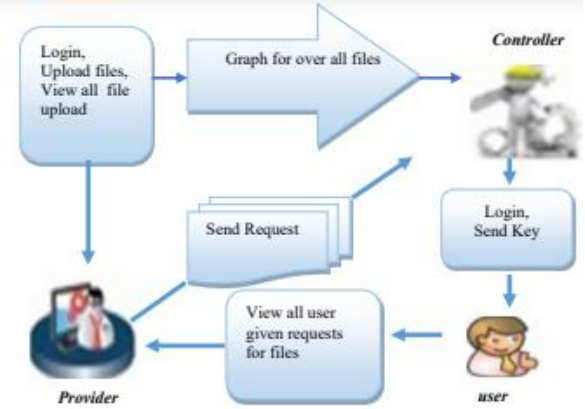

Figure 1 Architecture of Dynamic Load Balancer for Traffic Management

IV. SYSTEM ARCHITECTURE

Given below is an example implementation of a genetic algorithm in Java. Given a set of 3 datacenters, each datacenter can hold one binary value upto 15 which is the total capacity. The value is calculated as the number of 1s present in the data center. If there are 14 1s, then it is having maximum capacity. If there are no 1s, then it has the minimum capacity. This genetic algorithm tries to maximize the capacity function to provide a load consisting of the maximum capacity i.e. datacenter with 14 balance. This system has three modules namely the user, the provider and the controller.

1. User

User can Register, Login, View all files, send File Requests, View Notifications and Download File. User needs to firstly register by giving the information. After registration, user can login with username and watchword. User will read all files and provides request for transfer the file. User can view notifications and download the file.

2. Provider

Provider can Login, Upload files, View all uploaded file, View all requests, send Request to controller. Data provider needs to login with username and password. Provider accustomed transfer files and he/she will read all uploaded files. View all user given requests for files. Then data provider gives request to the controller.

3. Controller

Controller can Login, Send Key to user, Bandwidth optimization, Graph for over all requests, Graph for overall files. Data controller are used to control and monitor all data centers. The SDN management has been applied to alter the central control of the network, and to improve utilization of bandwidth for provider and to reduce the delay that occurs between end-users. Data controller needs to login with username and password. The controller can send key to user. The graph can also be generated depending upon the overall user request and can also graph can be generated on all the files that are uploaded.

V. RESULTS AND DISCUSSIONS

User could register and could login from a variety of local host. Since user does not have file access permission, they can get access from the provider. Hence the user sends request to the Provider. The user can Register, Login, View all files, send File Requests, View Notifications Download File User needs to firstly register by giving the information. After registration, user can login with username and watchword. User will read all files and provides request for transfer the file. The provider forwards the request to the controller and the controller sends back a key to the user. The user enters the key and downloads file.

The provider can also upload the files for the user to view and request It forwards the request to the controller as shown in Figure 2. The provider can Login and Upload files. It also views all uploaded file, view all requests, request to controller. The controller logs in and allocates the file to the user with maximum capacity as in Figure 3.

The controller accepts the request from the provider and sends key to the user. Controller Login Sends Key to user Bandwidth optimization Graph for over all requests Graph for overall files. Data controller used to control and monitor all data centers. The Central Control of the network is altered by applying SDN management by considering high bandwidth utilization for provider and low delay for users [7].

Data controller are accessed by login with username and password. View request from the info supplier so controller can send key to user. Generate graph based result on the overall user requests and all the uploaded files.

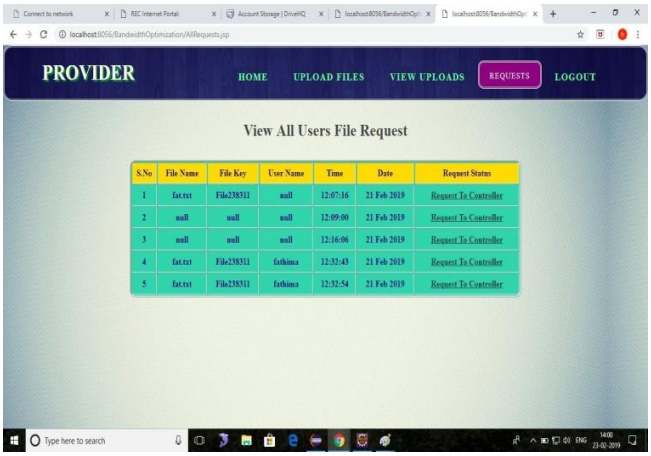

Figure 2 Provider requests to controller

Figure 3 Datacenter in the controller

The request allocation in geographically distributed datacenters is our main focus. The genetic logic is used to change the central control of the network by considering high bandwidth consumption for provider and low postponement for users. Specifically, the provider's demand of huge information measure utilization in the least datacenters and handlers' low delay necessities area unit each shapely supported the Ogden Nash talks game.

Then, the look of request allocation underneath those necessities is formulated as associate degree optimization downside, which is an integer optimization. To with efficiency solve such associate degree optimization downside introducing auxiliary variables for eliminating inequality constraints by proposing a request allocation algorithm, rather than directly put on the Logarithmic Smoothing technique. The existence and individuation of our best outcome and meeting of our rule are proved with the help of theoretical analysis.

VI. CONCLUSION

The dynamic load balancer for traffic management show that the rule will improve the efficiency of the information measure efficiency, consumption for the supplier and decrease the delay for users compared with greedy and section algorithm. As future work, a thorough study of the decentralized request allocation can be performed in each datacenter by using single controller.

REFERENCES

1. Amazon web services [Online]. Available: http://aws.amazon. com, 2014.
2. V. K. Adhikari, Y. Guo, F. Hao, M. Varvello, V. Hilt, M. Steiner, and Z.-L. Zhang, "Unreeling netflix: Understanding and rising multi-CDN moving picture delivery," in Proc. IEEE Conf. Comput. Commun., 2012, pp. 1620–1628.
3. A. Cockcroft. (2011). Netflix in the cloud [Online]. Available: http://velocityconf.com/ velocity 2011/public/schedule/detail/ 17785
4. B. Wong and E. G. Sirer, "Closestnode.com: associate degree open access, scalable, shared geocast service for distributed systems," operative Syst. Rev., vol. 40, no. 1, pp. 62–64, 2006.
5. H. Ballani, P. Costa, T. Karagiannis, and A. Rowstron, "Towards predictable datacenter networks," in Proc. ACM SIGCOMM Conf., Toronto, ON, Canada, 2011, pp. 242–253.
6. N. Laoutaris, M. Sirivianos, X. Yang, and P.Rodriguez, "Interdatacenter bulk transfers with netstitcher," in Proc. ACM SIGCOMM Conf., Toronto, ON, Canada, 2011, pp. 74–85.
7. A. Singh, M. Korupolu, and D. Mohapatra, "Server-storage virtualization: Integration and cargo equalisation in information centers," in Proc. ACM/IEEE Conf. Supercomput., 2008, p. 53.
8. R. Buyya, R. Ranjan, and R. N. Calheiros, "Intercloud: Utilityoriented federation of cloud computing environments for scaling of application services," in Proc. 10th Int. Conf. Algorithms Archit. Parallel Process., 2010, pp. 13–31.
9. A. Qureshi, R. Weber, H. Balakrishnan, J. V. Guttag, and B. M. Maggs, "Cutting the electrical bill for internet-scale systems," in Proc. ACM SIGCOMM Conf., 2009, pp. 123–134.
10. [10] J. Jinu Sophia, R .Angelin Hannah, J. Daisy, K. Jebima Jessy "Reducing The Duplication Data For Secure Authorization In Hybrid Cloud" in International Journal Of Advanced Research Trends In Engineering And Technology, Volume 3, Special Issue 19, April 2016.

AUTHORS PROFILE

S. Prithi is an Assistant Professor (SG) in the Department of Computer Science and Engineering at Rajalakshmi Engineering College, Chennai since 2018. She is currently pursuing her Ph.D in Anna University, Chennai in the Faculty of Information and Communication Engineering. She has got over 14 years of teaching experience in reputed engineering colleges in Coimbatore and Chennai. She completed her Masters in Computer Science and Engineering from Dr.M.G.R University, Chennai and her Bachelor's degree in Information Technology from CSI College of Engineering, The Nilgiris. Her research interest includes automata theory, machine learning, computational intelligence techniques and network security. She has published many papers in leading journals and conference proceedings.

Roxanna Samuel has completed her Bachelors in Computer Science and Engineering at Madras University in 1999 and Masters in Computer Science and Engineering at Sathyabama Deemed University in 2004. She is presently working as Assistant Professor (SG) in the Department of Computer Science and Engineering at Rajalakshmi Engineering College, Chennai. She has published many research papers in various journals and conferences covering several domains like, Big data, IoT, Data Mining. She has more than 12 years of teaching experience in various Engineering Colleges. Her area of interest includes Networks, IoT, Digital Technology, Data Analytics, Cryptography and Network Security. She has guided and is guiding students in both UG and PG Level and received the best faculty award in Computer Science and Engineering.

S. Ponmani received her B.E. degree from Madurai Kamarajar University, Tamil Nadu, India, in 2000 and M.E. degree from Anna University, India, in 2011. Currently, she is working as an assistant professor in Rajalakshmi Engineering College (Autonomous), Chennai, India. During her 18 years of experience she occupied different positions in academic and administration. Her research interest is Data mining. She is also interested in developing new teaching methodologies in the class room. Her education related papers are published in International Journals and have been selected in National and International Conferences.

J. Jinu Sophia is currently working as Assistant Professor in Rajalakshmi Engineering College, Chennai, India. She received the B.E degree in Computer Science from Anna University in 2006, M.E degree in Computer Science from Anna University in 2011 and currently Pursuing Ph.D degree in Sathyabama University, Chennai, India. Her research interests include Deep Learning, Natural Language Processing, Cloud Computing and Networking.