Load Balancing Using Joint Optimization of Bandwidth

Sunil Prakash J[1], Rikhil G R[2], Alen Mammen Ivan[3], Devender Singh[4], Vidhya Vani A[5]

[1] *student, department of computer science and engineering, SRM IST Ramapuram, tamilnadu, India*
[2] *student, department of computer science and engineering, SRM IST Ramapuram, tamilnadu, India*
[3] *student, department of computer science and engineering, SRM IST Ramapuram, tamilnadu, India*
[4] *student, department of computer science and engineering, SRM IST Ramapuram, tamilnadu, India*
[5] *asst professor, department of computer science and engineering, SRM IST Ramapuram, tamilnadu, India*

ABSTRACT

In large-scale Internet applications running on geographically distributed datacenters, such as video streaming, it is important to efficiently allocate requests among datacenters. To the best of our knowledge, existing approaches, however, either solely focus on minimizing total cost for provider, or guaranteeing QoS for end-users. In this paper, we apply the software defined network (SDN) controller to enable the central control of the entire network, and propose a joint optimization model to consider high bandwidth utilization for provider and low delay for users. We present the Nash bargaining solution (NBS) based method to model both requirements of provider's high bandwidth utilization and end-users' low delay. Specifically, we formulate the design of request allocation under those requirements as an optimization problem, which is NP-hard. To solve such hard optimization problem, we develop an efficient algorithm blending the advantages of Logarithmic Smoothing technique and the auxiliary variable method. According to the theoretical analysis, we verify the existence and uniqueness of our solution and the convergence of our algorithm. We conduct a large amount of experiments based on real-world workload traces and demonstrate the efficiency of our algorithm compared to both greedy and locality algorithms.

Keyword: - *Large-Scale Internet Applications , Datacenters , Software Defined Network , Joint Optimization Model , Bandwidth , NP-hard , Logarithmic Smoothing Technique , Nash bargaining solution .*

---

1. INTRODUCTION

Cloud Computing is a technology which depends on sharing of computing resources than having local servers or personal devices to handle the applications. In Cloud Computing, the word "Cloud" means "The Internet", so Cloud Computing means a type of computing in which services are delivered through the Internet. The goal of Cloud Computing is to make use of increasing computing power to execute millions of instructions per second. Cloud Computing uses networks of a large group of servers with specialized connections to distribute data processing among the servers. Instead of installing a software suite for each computer, this technology requires to install a single software in each computer that allows users to log into a Web-based service and which also hosts all the programs required by the user. There's a significant workload shift, in a cloud computing system.

Local computers no longer have to take the entire burden when it comes to running applications. Cloud computing technology is being used to minimize the usage cost of computing resources. The cloud network, consisting of a network of computers, handles the load instead. The cost of software and hardware on the user end decreases. The only thing that must be done at the user's end is to run the cloud interface software to connect to the cloud. Cloud Computing

Consists of a front end and back end. The front end includes the user's computer and software required to access the cloud network. Back end consists of various computers, servers and database systems that create the cloud. The user can access applications in the cloud network from anywhere by connecting to the cloud using the Internet. Some of the real time applications which use Cloud Computing are Gmail, Google Calendar, Google Docs and Dropbox etc.

1.1 LITERATURE SURVEY

Y. Guo, F. Hao et al., [1] proposed Netflix is the leading provider of on-demand Internet video streaming in the US and Canada, accounting for 29.7% of the peak downstream traffic in US. Understanding the Netflix architecture and its performance can shed light on how to best optimize its design as well as on the design of similar on-demand streaming services.

B. Wong and E. G. Sirer. et.al. [2] described ClosestNode.com is an accurate, scalable, and backwards-compatible service for mapping clients to a nearby server. It provides a DNS interface by which unmodied clients can look up a service name, and get the IP address of the closest server.

H. Ballani, P. Costa, T. Karagiannis, et al., [3] The shared nature of the network in today multi-tenant datacenters implies that network performance for tenants can vary significantly. This applies to both production datacenters and cloud environments. Network performance variability hurts application performance which makes tenant costs unpredictable and causes provider revenue loss.

N. Laoutaris, M. Sirivianos et al., [4] Large datacenter operators with sites at multiple locations dimension their key resources according to the peak demand of the geographic area that each site covers. The demand of specific areas follows diurnal patterns with high peak to valley ratios that result in average utilization across a day.

A. Greenberg, J. Hamilton et al., [5] the data centers used to create cloud services represent a signifi- cant investment in capital outlay and ongoing costs. Accordingly, we first examine the costs of cloud data centers today.

2. PROPOSED SYSTEM

We focus on the emerging request allocation problem in geographically distributed datacenters, and propose a joint optimization model to consider high bandwidth utilization for provider and low delay for end-users. Specifically, we present Nash bargaining solution (NBS) based method to model both the requirement of provider's high bandwidth utilization at all datacenters and end-users' low delay. We formulate the request allocation under those requirements as an optimization problem. Such optimization can be NP-hard. To solve it, we propose an efficient request allocation algorithm by introducing the auxiliary variable method to eliminate inequality, rather than directly applying the Logarithmic Smoothing technique. We perform theoretical analysis to prove the existence and uniqueness of our solution, and the convergence of our algorithm.

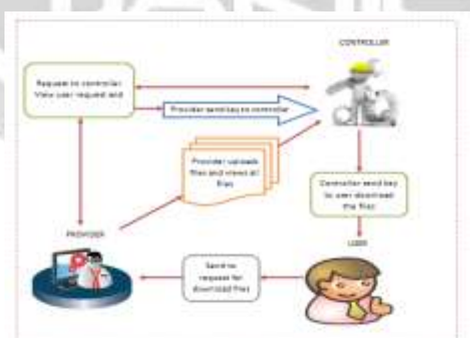

Fig -1: **System Architecture**

3. IMPLEMENTATION

User:
1. Register
2. Login
3. View all files

4. File Requests
5. View Notifications
6. Download File

User needs to firstly register by giving the information. After registration, user will login with username and password. User can view all files and give request for download the file. User can view notifications and download the file.

Data provider:
1. Login
2. Upload files
3. View all uploaded file
4. View all requests
5. Request to controller

Data provider needs to login with username and password. Provider used to upload files and he/she can view all uploaded files. View all user given requests for files. Then data provider give request to the controller.

Data controller:
1. Login
2. Send Key to user
3. Bandwidth optimization
4. Graph for over all requests
5. Graph for overall files

Data controller used to control and monitor all data centers. We apply the SDN controller to enable the central control of the network, and jointly consider high bandwidth utilization for provider and low delay for users. Data controller needs to login with username and password. View request from the data provider and then controller will send key to user. Generate graph based on the overall user requests and generate graph based on the all the uploaded files.

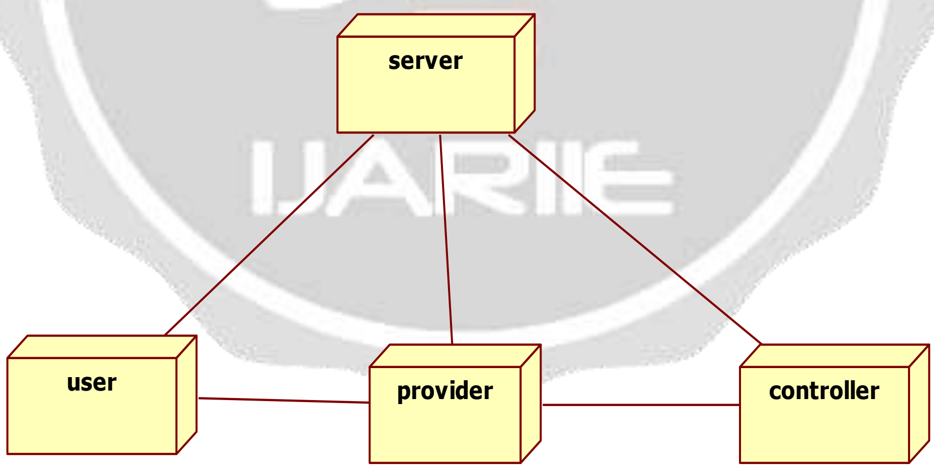

Fig -2: **Module Description**

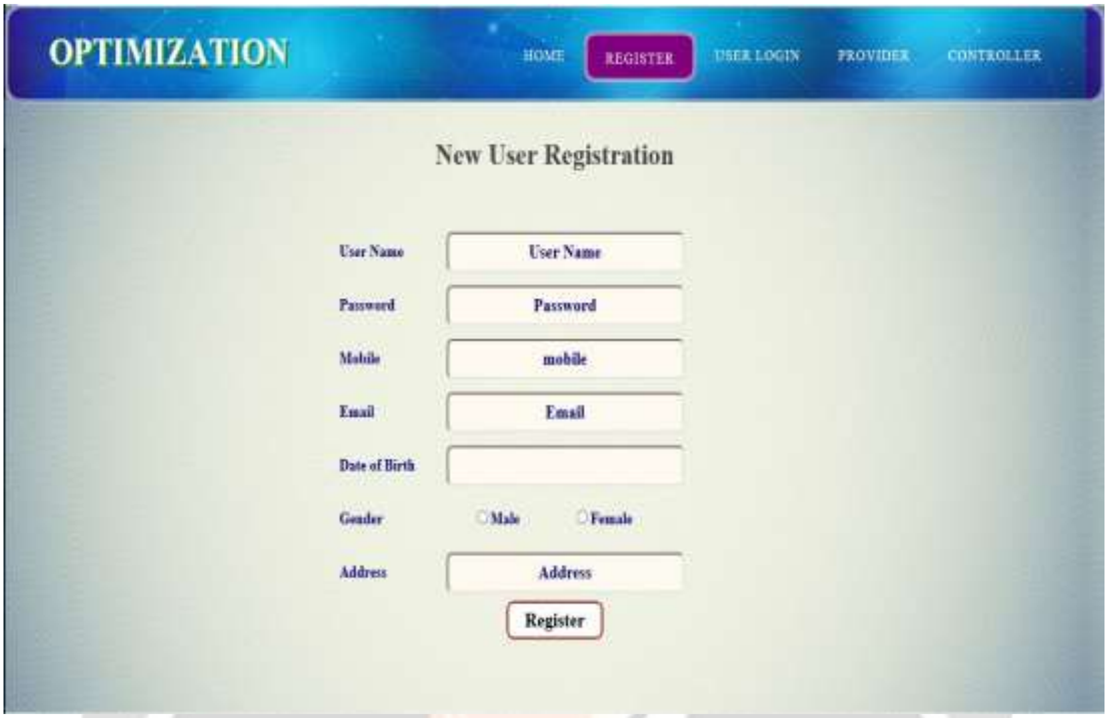

Fig -3: **Registration Page**

Fig -4: **login page**

Fig -5: **person details**

Fig -6: **list of files**

4. CONCLUSIONS

Our focus in this paper is the request allocation in geographically distributed datacenters. To efficiently allocate requests, we apply the SDN controller to enable the central control of the network, and jointly consider high bandwidth utilization for provider and low delay for users. Specifically, the provider's requirement of high bandwidth utilization at all datacenters and users' low delay requirements are both modeled based on the Nash bargaining game. Then, we formulate the design of request allocation under those requirements as an optimization problem, which is an integer optimization as well as NP-hard.

To efficiently solve such an optimization problem, we propose a request allocation algorithm by introducing auxiliary variables to eliminate inequality constraints, rather than directly applying the Logarithmic Smoothing technique. Theoretical analysis proves the existence and uniqueness of our optimal solution and the convergence of our algorithm. We empirically evaluate our algorithm based on real-world workload traces. The experimental results show that our algorithm can efficiently improve the bandwidth utilization for the provider and reduce the delay for users, compared with both greedy and locality algorithms.

5. 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 improving multi-CDN movie 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: An open access, scalable, shared geocast service for distributed systems," Operating 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.