Traffic Analysis of Content Delivery Network

Received: 02-Dec-2020 | Accepted: 18-Jan-2021

Durr E Najaf, Pinial K Butt, Arifa Bhutto

Abstract - Content delivery network (CDN's) have been developed to overcome the fundamental limitations of internet in terms of Quality of Service (QoS). A CDN duplicates the content of origin server to surrogate servers placed across the globe to deliver the contents to the end users in an efficient way. Content delivery on the web has received considerable research attention and the idea is to suggest an efficient networking infrastructure and replica management scenario through performance analysis of CDN. Despite its inherent limitations, even a not very complex infrastructure can show many characteristics about its ability to serve website and media traffic in a low latency model. Server utilization issues, request failures as well as mean response time are analyzed by testing different CDN policies. An efficient infrastructure is then suggested to provide guaranteed QoS for web contents using content delivery network. From this paper the readers can learn about four scenarios of CDN which are closest surrogate, Random surrogate, Load balancing surrogate and closest origin server. The scenarios are examined on the simulator od CDN where the virtual traffic passed from them to observe their response time, aborted requests and server utilization stats, with the help of generated data from the simulation results the results have been concluded.

Inspect Classification: C5620

Keywords: Content Delivery Network, Internet traffic, Cache proxy, Surrogate servers, Internet service providers.

I. INTRODUCTION

Internet traffic is increasing day by day, people use the Internet like resources hungry applications such as web objects (text, graphics and scripts), download-able objects (media files, software, and documents), applications (e-commerce, portals), live streaming media, on-demand streaming media, and social media networks [16].

Whenever user sends a request, sometimes it takes a lot of time to process or clients request are aborted due to network congestion, missing cache location or narrow bandwidth of the WLAN links. It is very much difficult to manage and deliver data through single main server. In order to store and deliver large amount of web contents, an efficient network infrastructure is required to provide service to users geographically at the edges of the globe [11].

The architecture of CDN is based on surrogate's servers that are distributed geographically at the edges of the globe [12]. Through CDN distribution node and allowing content providers to upload their data over these surrogate's servers. If content is not found at any surrogate server, the request is redirected to the other surrogate server [1].CDN reduces the hop to hop delivery and less number of hops mean more efficient network, less bandwidth utilization and reliable [9].

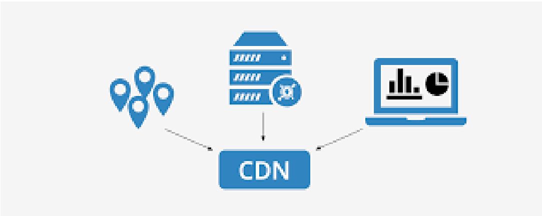

Figure 1: CDN Architecture

typical customers of CDN and wants to deliver their content to the end users in a very reliable manner.

A. Caching Proxy:

For narrow bandwidth users there is deployment of caching proxy by ISP. In order to improve performance and less bandwidth utilization caching proxies are deployed near to the end users, send request through these caches rather than sending to the origin server [2].The user browsing session goes through specific caching proxy when the entire configuration is properly done [8]. Different level of arrangement may be deployed by ISP such as local, regional, international referred as hierarchal caching as shown in Figure 2. This will improve performance and saved bandwidth.

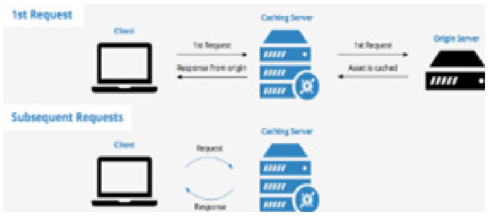

Figure 2: Cache Proxy Process

II. AIM OF THIS RESEARCH

A user sends request to access web content. It depends on infrastructure and utilization policy that client gets their

Durr E Najaf Dept: Information Technology Centre Sindh Agriculture University Tandojam, Pakistan

Pinial K Butt Dept: Information Technology Centre Sindh Agriculture University Tandojam, Pakistan

Arifa Bhutto Dept: IICT, University of Sindh Jamshoro Pakistan

request to be completed or not so it took some time. These challenges depends upon the network architecture and how the services are provided to the users which is basically a hop to hop delivery [4]. A server is placed at far distance then request will take to many hops to access control. In order to optimize how many hops it will take by a single request we can use trace-route

Network for delivering content. Refer to theFigure 1 for basic architecture of CDN. Media Internet advertisement companies, datacentres, command (tracert) to an internet address such as www.gmail.com.Servers located at far distance may result badly in response time or delays or poor service quality and it also depends on the congestion of web traffic [3]. In this paper the analysis of server utilization issues, request failures and mean response time of the network by using various content delivery policies [13]. In order to optimize how many hops it will take by a single request from the tracert command to an internet address. Estimated number of hops are shown in the figure 3.

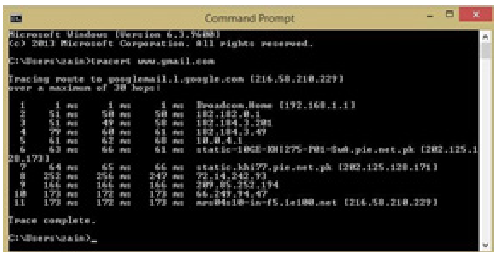

Figure 3: Hop to hop delivery

To achieve the results we design and simulate the architecture of CDN and analysis the quality of service for response time, aborted requests and server utilization.

III. RESEARCH METHODOLOGY

A. CDN Sim (simulator):

The simulation framework CDN Sim has been designed to provide real-time simulation for content delivery network (CDN), simulating the surrogate servers the TCP/IP protocol, and the primary CDN functionalities. The fundamental advantages of this simulating tool are its high throughput, its extensibility, and it's UI, which is use to configure its parameters [5]. CDN Sim gives a mechanized environment to directing experiment tasks and expressing client, server, and network statistics [15]. In CDN simulator the simulation container called bottle. Steps of the creation of bottles shown in figure 4.

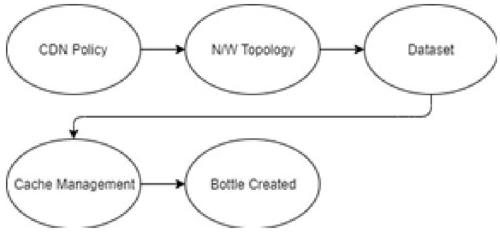

Figure 4: Process of bottle creation

B. SCENARIOS:

To perform experiments and get results we have 4 scenarios which are closest surrogate, random surrogate, load balancing surrogate and closest origin. In the topology we have 15 routers interconnected to each other with the link speed of 10Mbps and one origin server. In all of these scenarios user send request to access some data, user will direct to the surrogate servers. In the case of cache miss user will redirect to the other nearest possible surrogate servers [6] [7]. In the load balancing scenario user will only redirect to another server if the limit reaches to 95% in order to balance the load [10]. Other scenarios will work same. The graphs below shows the results of response time, aborted requests and server utilization on 100 constant clients and 30surrogate servers for these four scenarios [14].

(1). CLOSEST SURROGATE:

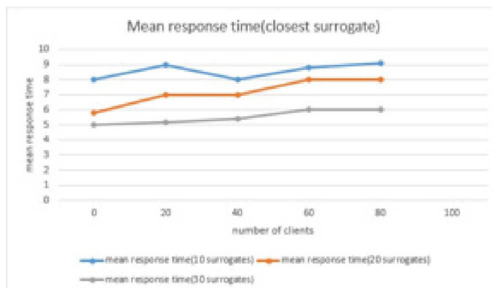

Figure 5:Response Time (Closest Surrogate)

Description: Least response time is 5 seconds.

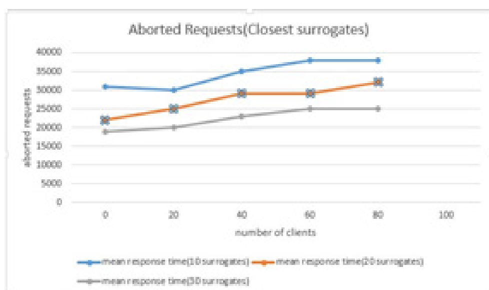

Figure 6:Aborted Requests (Closest surrogate)

Description: Least aborted requests are 1800.

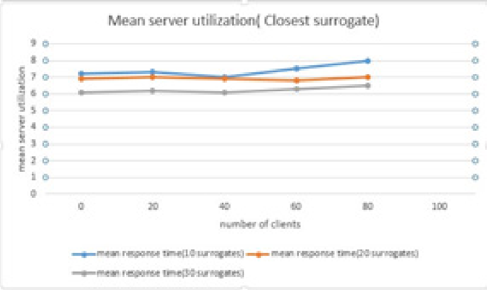

Figure 7: Server Utilization (Closest Surrogate)
Description: Least server utilization is 0.6%.

(2). RANDOM SURROGATE:

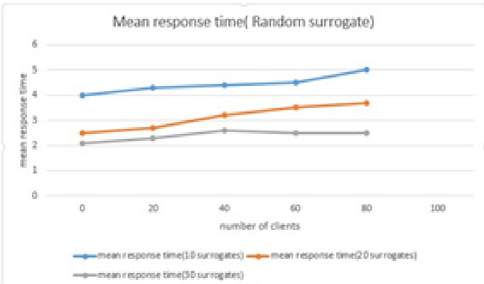

Figure 8: Response time (Random Surrogate)
Description: Least response time is 4 seconds

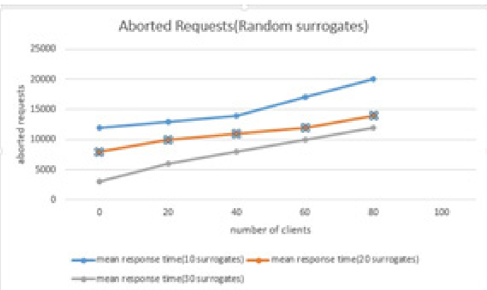

Figure 9: Aborted Requests (Random Surrogate)
Description: 0 aborted requests.

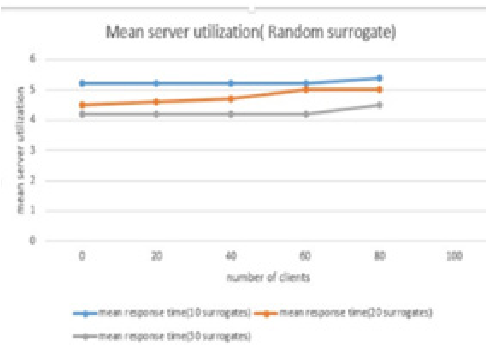

Figure 10:: Server Utilization(Random Surrogate)
Description: Least server utilization is 0.4%.

(3). LOAD BALANCING SURROGATE:

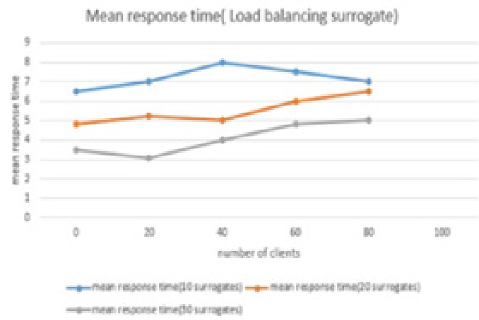

Figure 11: Response Time (Load Balancing Surrogate)
Description: Least response time is 3 to 4 seconds.

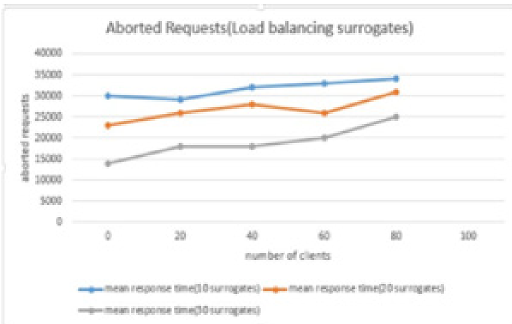

Figure 12: Aborted Requests (Load Balancing Surrogate)
Description: Least aborted requests are 14000.

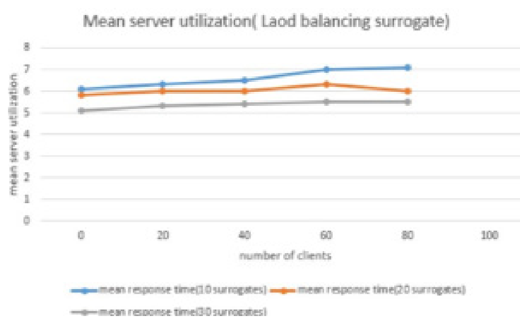

Figure 13: Server Utilization (Load Balancing Surrogate)
Description: Least server utilization is 0.5%.

(4) CLOSEST ORIGIN SERVER:

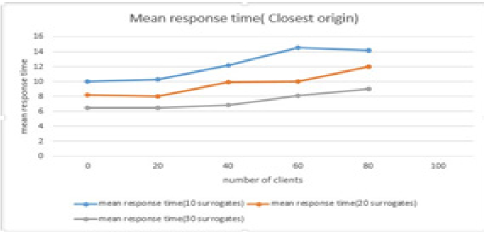

Figure 14: Response time (Closest Origin Server)

Description: Least response time is 6 to 8 seconds

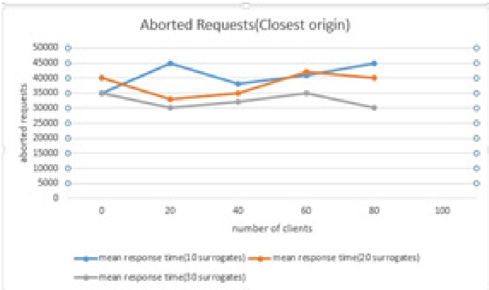

Figure 15: Aborted Requests (Closest Origin Server)

Description: Least aborted requests are 40000.

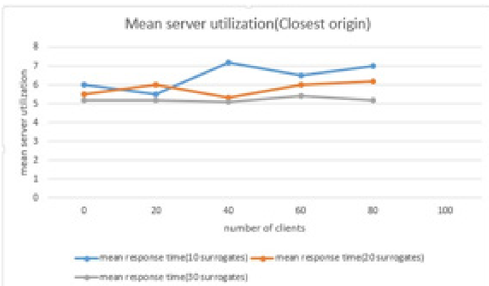

Figure 16: Server Utilization (Closest Origin Server)

Description: Least server utilization is 0.5%.

IV. RESULTS

After getting results from the scenarios on the parameters of mean response time, aborted requests and mean server utilization for 100 constant clients and 30 active surrogate servers here are the comparison result decides which scenario is best for CDN.

A. COMPARISON OF MEAN RESPONSETIME:

The comparison of mean response time for four different scenarios is shown in Figure 17 by taking 100 clients with varying number of surrogates. The mean response time for four different policies by taking 30 constant Surrogates but varying number of clients is shown in Figure 18.

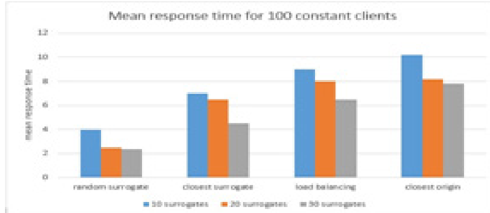

Figure 17: Response Time For 100 Constant Clients

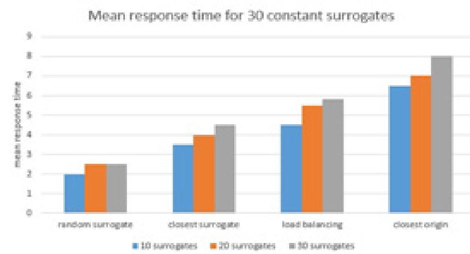

Figure 18: Response Time For 30 Constant Surrogates

Description: Least response time is found in random surrogate and then closest surrogate.

B. COMPARISON OF MEAN SERVERUTILIZATION:

The comparison of mean server utilization for four different scenarios is shown in Fig19 by taking 100 clients with varying number of surrogates. The mean server utilization for four different policies by taking 30 constant surrogates but varying number of clients is shown in Figure 20.

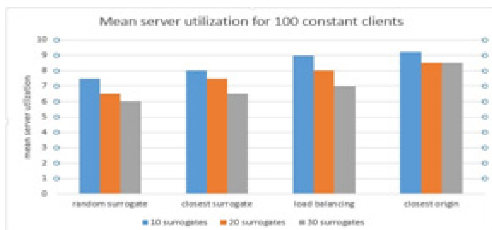

Figure 19:Server Utilization For 100 Constant Clients

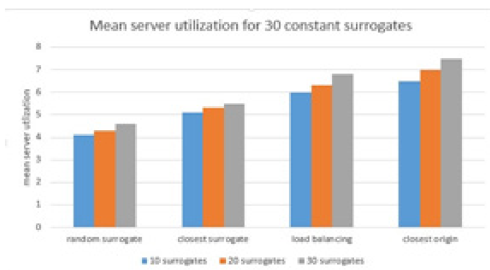

Figure 20: Server Utilization for 30 Constant Surrogates

Description: Least server utilization is found in random surrogate then closest surrogate.

C. COMPARISON OF ABORTED REQUESTS:

The comparison of aborted requests for four different scenarios is shown in Figure 21 by taking 100 clients with varying number of surrogates. The aborted requests for four different policies by taking 30 constant surrogates but varying number of clients is shown in Figure 22.

Figure 21: Aborted Requests for 100 Constant Clients

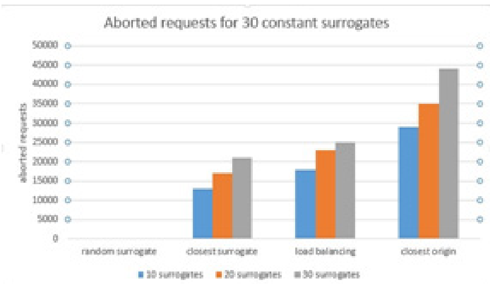

Figure 22: Aborted Requests For 30 Constant Surrogates

Description: Least aborted requests are found in random surrogate scenario then closest surrogate scenario.

V. CONCLUSION

Existing infrastructure of CDNs have always been evolved to provide efficient mechanism for delivery of content to the clients with less jitter, no delays and with better utilization of existing bandwidth. In our work, we have analysed the performance of CDNs using the network simulation software CDN sim, and many experiments have been performed based on four different scenarios. The experimental setup are performed by considering the specific number of clients per surrogate server and as a result mean response time, number of aborted requests and mean server utilization are found to be different for different scenario. The best response are found to be in closest surrogate and random surrogates scenarios with least mean response time, aborted requests and low server utilization to deliver specific contents to the clients. The preferable scenario is closest surrogate.

Moreover, issue of bandwidth utilization is solved by setting a suitable number of surrogates for local area. Although CDN provide dominance, there is always a chance for constant development. The idea is to develop a simulation framework for content delivery network based on new libraries to support new trends such as depository, scalability of active content, applicability of any casting, dynamic content usage, flexible for content delivery network.

REFERENCES

[1] Andrea Bianco, Reza Mashayekhi, Michela Meo.”On the energy consumption in content delivery network”. 2210-5379 Elsevier Inc.2017.

[2] Andrea Bianco, Reza Mashayekhi,Michela Meo. “Energy Consumption for Data Distribution in Content Delivery Networks”. Green Communications Systems and Networks 978-1-4799-66646/16/$31.00 2016 IEEE.

[3] Ali A. Haghhighi, Shahram Shah Haidri, Shahram Shahbaz Panahi, “Dynamic QoS Aware Resource Assignment in Cloud based Content-Delivery Networks” ,IEEE 2017.

[4] Bo Bai, Li Wang, Zhu han, Wei Chen, and Tommy SvenSSon. “Caching based socially-aware d2d Communications in wireless Content delivery networks: a hypergraph Framework”. 1536-1284/16/$25.00 © 2016 IEEE Wireless Communications • August 2016.

[5] Daniel S. Berger, Ramesh K. Sitaraman, Mor Harchol-Balter. “Adapt Size: Orchestrating the Hot Object Memory Cache in a Content Delivery Network”. ISBN 978-1-931971-379 2017.

[6] Francisco Elias Rangel-Patiño, José LuisChávez-Hurtado, Andres Viveros-Wacher,José Ernesto Rayas-Sánchez, and NagibHakim. “System Margining Surrogate-Based Optimization in Post-Silicon Validation”.0018-9480 © 2017 IEEE.

[7] Ghazaleh Eslami Abolfazl Toroghi Haghighat. “A new surrogate placement algorithm for cloud-based content delivery networks”. © Springer Science Business Media, LLC 2017.

[8] Jihoon Sung, Minseok Kim, Kyongchun Lim, and June-Koo Kevin Rhee. “Efficient Cache Placement Strategy in Two-Tier Wireless Content Delivery Network”.15209210 (c) 2016 IEEE.

[9] Liang Zhou, Dan Wu, Zhenjiang Dong, and Xuelong Li. “When Collaboration Hugs Intelligence: Content Delivery over Ultra Dense Networks”. 0163-6804/17/$25.00© 2017 IEEE.

[10] M. Zubair Shafiq, Amir R. Khakpour, Alex X. Liu.

"Characterizing Caching Workload of a Large Commercial Content Delivery Network". 978-1-4673-9953-1/16/$31.00 ©2016 IEEE.

[11] Nicolas Herbaut, Daniel Negru, Yiping Chen Pantelis A. Frangoudis, Adlen Ksentini."Content Delivery Networks as a Virtual Network Function: a Win-Win ISP-CDN Collaboration". 978-1-5090-13289/16/$31.00©2016 IEEE.

[12] Narjes Taghigh Jahromi, Roch H. Glitho, Adel Larabi, Richard Burner. "An NFV and Micro service Based Architecture for On the Fly Component Provisioning in Content Delivery Networks". 978-1-5386-47905/18/$31.00 ©2018 IEEE.

[13] Sara Retal, Miloud Bagaa, Tarik Taleb and Hannu Flinck. "Content Delivery Network Slicing: QoE and Cost Awareness". 978-14673-8999-0/17/$31.00 ©2017 IEEE.

[14] Saif Ul Islam, Hasan Ali Khattak, Faisal Fayyaz Qureshi, Victor Chang, Jean Marc Pierson. "Leveraging Utilization as Performance Metric for CDN enabled Internet of Things". S0263-2241(19)306712, 9 July2019.

[15] Sayan Sen Sarma, S. K. Setua "Uniform load sharing on a hierarchical content delivery network interconnection model". DOI 10.1007/s11334-016-0279-52016.

[16] Zheng wang, Jun Huang, Scott Rose "Evolution and challenges of DNS-Based CDNs". S2352-8648(17)30073-1 2017.