

Object Caching Scheme to Render Reduction of Latency for Webpages

^{*1}Satti Sairam Reddy, ²P. Malathi

^{*1}UG Student, ²Assistant Professor, Saveetha School of Engineering,
Saveetha Institute of Medical and Technical Sciences, Chennai, India

^{*1}sattisairamreddy@gmail.com, ²maludhurkka@gmail.com

Article Info

Volume 83

Page Number: 1699 - 1702

Publication Issue:

March - April 2020

Article History

Article Received: 24 July 2019

Revised: 12 September 2019

Accepted: 15 February 2020

Publication: 15 March 2020

Abstract

Web pages are related to web services and stand as important research topic in today's world. The number of critical objects in the web pages along with latency reduction is also one of the parameters in case of web services. Our work is focused towards identifying these objects that are said to be critical and also store them in cache to decrease the time spent in rendering. The parameters are studied in the CDN (Content Delivery Network) provider side and also only limited objects are stored in cache to retrieve faster. The optimization problem said to be one-shot can be proved to NP-hard. Our proposed work continues to develop greedy algorithm with effective computation time. Simulation results show that our system reduces the latency of the system as well as traffic produced by the network which in turn increases the hit ratio of the cache. This means the miss ratio of the cache is obviously higher in the proposed system.

Keywords: Content Delivery Network, Latency, Optimization, Web pages

1. Introduction

Services including multimedia contents such as streaming videos as well as social networks like facebook and twitter can be used in designing web pages. Objects in multimedia to be included in web pages are text, image as well as video. The amount of web pages is increasing nowadays so that it becomes useful to the users. For loading the webpage, time for loading is the main parameter to be noted that indicates the usage of users. Thus the latency produced by web pages need to be decreased which plays a major role in the industry of networking.

In case of reducing latencies, many issues arise. The first issue to be focused on objects including both static and dynamic like images that may be stored on various servers and also have compound dependencies. Subsequently, to download a particular object, the activity depends on other objects too. Another

issue to be focused on is the contradiction between users and the providers of CDN. The former prefers latency of short length and the latter concentrates on the object delivery optimization. The webpages of modern systems have different impact in case of latencies of webpage.

Our proposed system mainly focus on reducing the latency of webpages and thereby network traffic gets reduced. The paper is organized as follows: Section 2 presents the literature review, Section 3 discusses the proposed system, Section 4 presents the results obtained from the proposed work and Section 5 concludes the paper.

2. Literature Review

To study the performance of caching various techniques have been developed due to the complexity of web pages. The webpages study brought our many findings including

the dynamic content [1]. Several objects both in static or dynamic state is the main composition of webpages also including dependencies in complex level. The literature work can be further classified as optimizing the caching of CDN and also as web loading on client side.

Optimizing the caching of CDN: Algorithms used for web caching are several in number available in the literature [2] [3]. Several schemes are based on either LFU or LRU [4] [5]. Placing objects in the caching system in hierarchical manner to reduce the latencies under conditions like cache bandwidth are also studied [6] [7]. Algorithm named Greedy Dual Size identifies the significance of web page objects which reduces the latency of webpage finally. In the efforts of optimizing the cache scheme, resources related to objects and networks are studied [8].

Web loading on Client side: The loading time of web pages from the side of client is being studies by many researchers [9]. Mainly the browsers that function as proxy offload the work of computation on to the cloud thereby lessens the client side bandwidth [10]. Object level web content prioritizing remained as the major focus of research on client side [11].

Generally a webpage iscomprising of different multimedia objects. It shows different video streaming, social networks and E commerce. While loading the webpage it is the key metric which indicates experiences and business revenues[12][13]. By the given reports it shows that the shopzilla increased its revenue by 12% as the load time also reduced to 6 to 1.2seconds. Here the problem that caused by the reducing webpage latencies is the attracted efforts from many industries[14].At present the large scale of the adoption of geo distributed contents(CDNs).

By reducing the challenges of webpage latencies there are two aspects [15]. The first aspect says that it contains of hundreds of static and dynamic objects whereas, the javascript(JS), images, cascading style sheets (CSS) [16] [17].They exhibit complex dependencies. Downloading of these things results in preceding, delay of these objects happens something relates to the objects. The second aspectsshow the contradiction between the users and CDN providers [18] [19]. Users have a very less latency, minimizes the bandwidth costs [20].Webpage have a similar popularity level. Whiledownloading it renders [21].

3. Proposed Method

Caching objects over content delivery network is the important field in research area of web services. The

providers sources consists of several services which in turn is made up of various objects. Source of contents helps in replication of objects to CDN nodes that helps the end-users in reducing the latency. When a service is requested by the user, these request are directed towards CDN nodes that appear to be closer geographically. From there, they are directed to several other nodes in case of miss in the cache for that particular object. After getting back either all the objects or the object parts, uninterruptable services is provided to users. Greedy algorithm is used in CDN nodes to reduce the latency of web pages to provide effective web services.

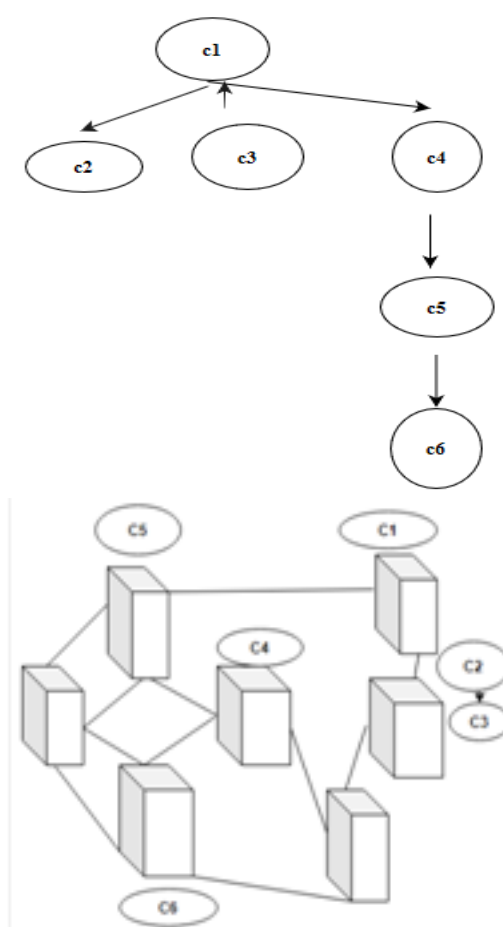


Figure 1: Content Delivery Network Architecture

4. Methodologies

To demonstrate the effect of cdn cache on webpage stacking we lead a genuine estimation.

In the estimation experiment we creep websites and record the page loading occasions as http file.

HAR is ajson formatted chronicle record position for logging of an internet browser cooperation with a web page, eg http demands and relating reactions and page on load time. Home pages of akami clients are picks as the

main sites .akamai gives pragma headers to examine orinvestigate akamai content conveyance. At the point when the pragma header is set in a http demand header,relating troubleshoot Data will be joined to the http reaction header.

For instance when akamai-x-get-demand id is set in he pragma header. X-AKAMAI-REQUEST-ID" will return in the http reaction header with a spot isolated rundown of CDNs

5. Results and Discussion

The topology of Content Delivery Network is simulated with many nodes and selection of nodes is done separately. To extract all objects including object attributes crawled webpages were suggested. In case of requests, the nodes arrive at the same time because of the factor to balance the load. The comparison graph showing the execution time of Algorithm 1 and Algorithm 2 is shown is the Fig.2. This shows the efficacy of the proposed system with Algorithm 2.

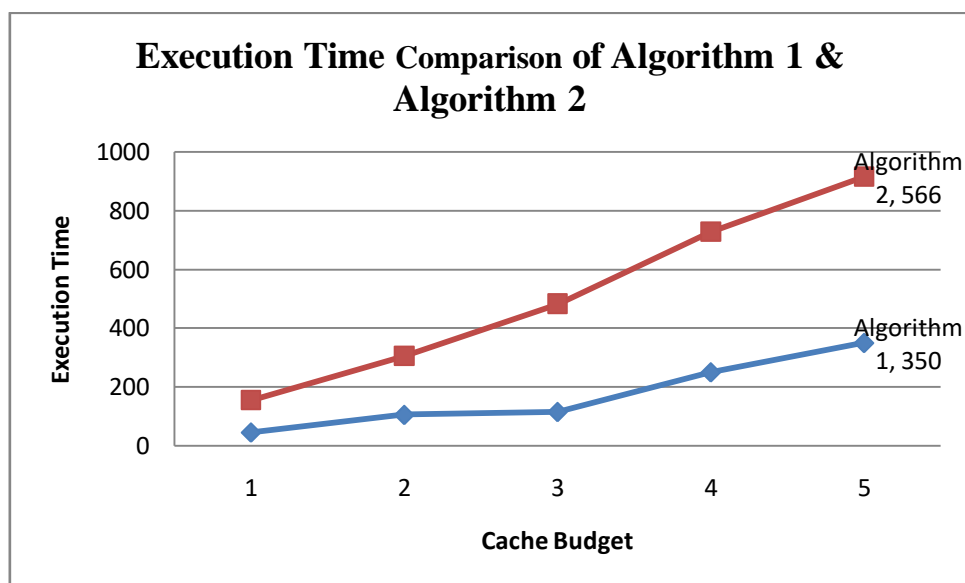


Figure 2: Comparison of Execution Time

6. Conclusion

Users using the web pages mainly focus on reducing the latency of web pages thereby the paper gives importance to introduction to caching scheme. Based on the reports of most providers of CDN, only limited objects will be stored in the cache memory and also the experimental results show that the critical objects were not cached which acts as an optimization problem that is constrained. Online algorithm developed in the proposed work reduces the rendering time in case of webpage usage which also increases the hit ratio of cache thereby decreasing the traffic created in the network.

References

- [1] X. Gu and K. Nahrstedt, "Distributed multimedia service composition with statistical qos assurances," IEEE transactions on multimedia, vol. 8, no. 1, pp. 141–151, 2006.
- [2] J. Hamilton, "The cost of latency." [Online]. Available: <http://perspectives.mvdirona.com/2009/10/the-cost-of-latency/>
- [3] Shopzilla, "faster page load time = 12% revenue increase." [Online]. Available: <http://www.strangeloopnetworks.com/resources/infographics/web-performance-and-ecommerce/shopzilla-faster-pages-12-revenue-increase/>.
- [4] L. Zhou, "On data-driven delay estimation for media cloud," IEEE Transactions on Multimedia, vol. 18, no. 5, pp. 905–915, 2016.
- [5] "Qoe-driven delay announcement for cloud mobile media," IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 1, pp. 84–94, 2017.
- [6] Y. Wen, X. Zhu, J. J. Rodrigues, and C. W. Chen, "Cloud mobile media: Reflections and

- outlook,” IEEE Transactions on Multimedia, vol. 16, no. 4, pp. 885–902, 2014.
- [7] J. Tang, W. P. Tay, and Y. Wen, “Dynamic request redirection and elastic service scaling in cloud-centric media networks,” IEEE Transactions on Multimedia, vol. 16, no. 5, pp. 1434–1445, 2014.
- [8] Z. Wang, L. Sun, W. Zhu, S. Yang, H. Li, and D. Wu, “Joint social and content recommendation for user-generated videos in online social network,” IEEE Transactions on Multimedia, vol. 15, no. 3, pp. 698–709, 2013.
- [9] J. Wang, “A survey of web caching schemes for the internet,” ACM SIGCOMM Computer Communication Review, vol. 29, no. 5, pp. 36–46, 1999.
- [10] J. P. G. Ortiz, V. G. Ruiz, M. F. Lopez, and I. García, “Interactive transmission of jpeg2000 images using web proxy caching,” IEEE Transactions on Multimedia, vol. 10, no. 4, pp. 629–636, 2008.
- [11] M. Z. Shafiq, A. R. Khakpour, and A. X. Liu, “Characterizing caching workload of a large commercial content delivery network,” in INFOCOM’16, 2016, pp. 193–206.
- [12] Z. Wang, W. Zhu, M. Chen, L. Sun, and S. Yang, “Cpcdn: content delivery powered by context and user intelligence,” IEEE Transactions on Multimedia, vol. 17, no. 1, pp. 92–103, 2015.
- [13] S. Puzhavakath Narayanan, Y. S. Nam, A. Sivakumar, B. Chandrasekaran, B. Maggs, and S. Rao, “Reducing latency through pageaware management of web objects by content delivery networks,” in Proceedings of the 2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science. ACM, 2016, pp. 89–100.
- [14] V. Agababov, M. Buettner, V. Chudnovsky, M. Cogan, B. Greenstein, S. McDaniel, M. Piatek, C. Scott, M. Welsh, and B. Yin, “Flywheel: Google’s data compression proxy for the mobile web,” in 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15), 2015, pp. 367–380.
- [15] Google, “Spdy: An experimental protocol for a faster web.” [Online]. Available: <https://www.chromium.org/spdy/spdy-whitepaper>
- [16] M. Butkiewicz, D. Wang, Z. Wu, H. V. Madhyastha, and V. Sekar, “Klotski: Reprioritizing web content to improve user experience on mobile devices,” in 12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15), 2015, pp. 439–453.
- [17] I. Logic. [Online]. Available: <https://www.instartlogic.com/>
- [18] Z. Li, M. Zhang, Z. Zhu, Y. Chen, A. G. Greenberg, and Y.-M. Wang, “Webprophet: Automating performance prediction for web services.” in NSDI, vol. 10, 2010, pp. 143–158.
- [19] X. S. Wang, A. Balasubramanian, A. Krishnamurthy, and D. Wetherall, “Demystifying page load performance with wprof,” in Presented as part of the 10th USENIX Symposium on Networked Systems Design and Implementation (NSDI 13), 2013, pp. 473–485.
- [20] S. Souders, “Onload event and post-onload requests.” [Online]. Available: <https://www.stevesouders.com/blog/2013/02/26/reloading-post-onload-resources/>
- [21] R. Netravali, A. Goyal, J. Mickens, and H. Balakrishnan, “Polaris: faster page loads using fine-grained dependency tracking,” in 13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16), 2016.