International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-6 Issue-5, June 2017

Analysis and Advancement of Page Replacement Algorithm for web Proxy Server

Divyanshu Atre, Lalit Gehlod

Abstract: ***World Wide Web is growing rapidly and number of users is increasing day by day to access web pages. A Proxy server intercepts all client requests, and provides responses from its cache or forwards the request connects to the real server. Proxies were invented to add structure and encapsulations to distributed systems. The problems of web servers are heavy network traffic and Latency etc. To overcome these problems, proxy server caching is one of the solution. Proxy caching improve the speed of service requests by fetching the store web pages from an earlier request through the same point or even other point. In this research work Page Replacement Algorithm (LRU, LFU, FIFO) have analyzed for proxy server caching and proposed the randomly caching page replacement algorithm. It increases the hit rate with time and reduced execution time on proxy server cache. With the help of proposed algorithm, the performance of proxy server caching has been improved in terms of hit rate and time parameter.***

Index Terms: Proxy Server, Web cache, Latency, Page Replacement Algorithm (LRU, LFU, FIFO) etc.

I. INTRODUCTION

The internet or World Wide Web is a global network system which is interconnected to the different computer nodes. World Wide Web is a network of network which has many public, private government and academic networks. The internet users are increasing day by day. It has more than 3.4 billions of users . Due to the large number of users on internet, it is not possible to satisfy the demands of every user in an proficient manner. Users suffer from the slow network environment, weak response time and heavy server load etc. Therefore to overcome network load, low response time and heavy server load Proxy server is intermediated between the clients and the server. Caching technique allow organization to improve system performance. Web proxy cache has to play an important role to reduce the latency, heavy network traffic and heavy server load. As the cache size is limited, it is not possible to keep all the demanded web pages in a cache.

When the cache is already full and a request for web page appears, which is not present in the cache then a miss occurs. By the page replacement algorithms, one of the pages is to be eliminated to create space for new arrival page which is requested by clients. The better utilization and optimization of cache is done by the selection of good and efficient page replacement policy. Mostly used page replacement algorithms are LRU, LFU and FIFO etc [2,3]. The requests of Client's web pages are fulfilled by cache memory at proxy server. When client request for a web page, initially the demand request will go to proxy server. Proxy server will check the demand web page in its cache. If requested web page is found in Proxy server ,it will check the demand web page in its cache. If requested web page is not found then demand request will be forwarded to the main web server and the demand web page is redirect to the proxy server and stores the copy of requested demand page and served to clients.

Cache Page replacement algorithms play a central role in the response time reduction by selecting a subset of documents for caching so that a given web system performance is maximized. The aim of this paper is to improve web system performance through caching technique with design and implementation of a web proxy caching algorithm which can be increase hit rate and improve execution time . The remaining sections of this paper consist of three different parts. In related work section we focus on the work done in the performance improvement of proxy caching. The result and discussion section focus our finding and significant. In conclusion conclude our work with limitation and future scopefigures and tables.

II. RELATED WORK

Web proxy caching are huge active research areas. Network Traffic analysis of Web accessibility from the view of browsers proxies is most demanding research era previous tracing studies were bounded in number of requests, request rate and variety of population. The recent studies of tracing have wide scope and varied area. In addition to static analysis, some studies have also used to trace ambitious cache simulation to characterize the sharing property and localities of very large suggestions and to study the special effects of aborted connections, persistent connections and cookies on the concert of proxy caching. Many researcher have performed useful studies to improve current caching techniques in the field of proxy server. Recent studies on traffic of web are based on the analysis of Web accessibility from the approach of browsers proxies.

John Dilley [3] reported on the implementation and characterization of two newly proposed cache policies, LFU with Dynamic Aging (LFUDA) and Greedy Dual Size – Frequency (GDS-F) in the squid cache.**Ismail Ari [4]** proposed an adaptive Caching Scheme using Multiple Experts. It suggested that use of machine learning algorithms to choose the present best mixture of policies or policy by permitting each adaptive node to tune itself based on load of work.

Manuscript published on 30 June 2017.
***** Correspondence Author (s)
Mr. Divyanshu Atre, Department of Computer Engineering, Institute of Engineering and Technology, DAVV, Indore (M.P)-452017, India. E-mail: divyanshuatre126@rediff.com
Mr. Lalit Gehlod, Department of Computer Engineering, Institute of Engineering and Technology, DAVV, Indore (M.P)-452017, India.

Martin Arlitt [5] introduced virtual caches, an approach for getting better the performance of the cache for multiple metrics simulteneously.
J. Almeida [6] liked to further explore the performance of the new cache replacement policies under a more realistic proxy workload. The best way to do this is by using a benchmark able to generate a workload appropriate for a proxy, such as the wisconsin proxy benchmark. **Richa Gupta [7]** have proposed pair of replacement algorithms MFMR and AF-LRU for L1 and L2 cache for proxy server which performs better result than existing pair of replacement algorithms.**Irani [8]** have introduced greedy dual size, which have as a feature of locality with size and cost concerns in a simple and non parameterized fashion for elevated performance.. **Yong Zhen Guo [11]** proposed web page prefetching technique, they must be able to guess the next set of web pages that will be entréed by users and "Page-Rank Like Algorithm" was proposed for conducting web page prediction. **Golan [10]** proposed an optimal offline algorithm for page replacement in multi-level cache based on an algorithm for the rest listing updated problem and the DEMOTE operation. **Shiva Shankar Reddy [11]** have proposed a new method of caching HTTP Proxy servers which gets lower bandwidth by maintaining a web cache of internet objects. **Dr. Murali Bhaskaran [12]** discussed various data pre-processing techniques that are approved out at proxy server entrée log which generate web entrée pattern. These patterns are used for further application. **Song Jiang [13]** have proposed DULO algorithm, which utilizes both spatial and temporal locality in buffer cache management. **Ahcia Trivino-Cabrera [14]** performed a trace driven simulation study of the replacement policies has been devloped for the traffic generated by each considered content type. It had studied six replacement algorithms LFU, LFUDA and LRU, other three are specially developed for web documents GDSF, GD and GD Size, finally concluded that no replacement strategy do better than the other for all content type.

III. PROBLEMS WITH THE CACHING OF WEB PROXY

Web caching system should get better performance for network service providers, network administrator and end users, it can be accepted as an effectual way to grow the speed of web access rate, low server load, low latency perceived by the end users, improve response time and reduce network traffic to the end of users.

A. Load Maintaining

The situation occurs at any time for plentitude of clients who wants to wish simultaneously access content or access some services from its local cache with particular server. If the position is not provisioned to handle with these entire clients request simultaneously, service may be lost . Several ways to overcome this issue have proposed. The most frequently used method that is called caching. This caching technique saves the copy of popular web pages or services during the Internet, this spreads the job of serving a web page or service across different servers. Caching enhance the retrieval of web pages that access by same client or even other clients. It improves and increases the performance of server by reducing lot of server load.

B. Transparency

Transparency of cache systems allows users to get the advantages of caches without knowing that they exist, or without knowing to their location. The advantages of this technique are user friendly, no configuration detail is required by the end user and users can not escape the cache.

C. Scalability

It is very important that the cache system should be scalable as the servers and number of users increases day by day. It can be collective or stand-alone caches and clustered. Stand-alone caches are easy to maintain and handle the systems. However, cooperation between caches provided extra information about caching data, which could be communicated between caches without showing to the initilized server.

D. Cache Miss

Cache systems should be capable of fluently managing cache misses. When a cache miss occurs, a decision should be taken on where to be forwarded the request. Also a cache system should be decided that which data is to be cached or should all cached data be present equally.

Design and Implementation of Proposed Proxy Caching Algorithm

Web caching technique is a propitious technology in Web environment and in web caching, if the client requests a demand page from a server, it forwards to the server and will give response from the server. Which objects are cached form where location, web caching technology can be classified into following categories i.e., server-end proxy caching and client-end proxy caching.

If the request is validated by the filter, the proxy gives the resource by linking to the appropriate server and server serves the request service on the basis of the client's request. A proxy server may optionally change the server's response or the client's request and sometimes it may serve up the request without accessing to specified server. In this case, it caches response from its local cache

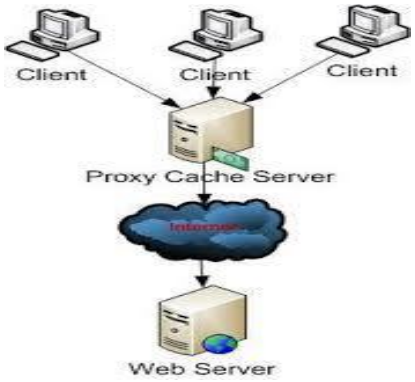

Figure 1 show a proxy server with cache memory which runs with a lot of characteristics such as reduces Latency time, load on web server, network traffic. This architecture also helps increase browsing of pages. In this system according to the proxy cache filled and new demand page request dismount at proxy a page replacement algorithm settle on which page has to be extrude from the cache. E

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

very hit in cache add the time counter to count the cache hit in millisecond. The page replacement algorithm is decided the efficiency of system.

Randomly Page Replacement Caching Algorithm

There has been embracing theoretical and pragmatic work done on exploring policies of web caching that achieve good result under different enumeration metrics. A lot of algorithms have proposed and initiate efficient for caching of web proxy. These algorithms collection from simple generic schemes such as least-recently used (LRU), least- frequently used (LFU), first-in first-out (FIFO) and various size-based algorithms, to compound hybrid algorithms such as LRU-Threshold, which bear a resemblance to LRU with limit of size on single cache elements, lowest relative value (LRV), which is used to size , cost and last position time to evaluate its utility, and Greedy Dual, which combines locality, cost and size consideration into this algorithm. Shows the proposed Randomly Page Replacement Caching Algorithm. This algorithm is focused on the characteristic i.e. algorithm of swapping documents with time counter.

Randomly Page Replacement Caching Algorithm (RPRCA)

If there is a page p is available in Cache:

1. Serve up first such p and does not mark the page
2. IF again p find than marks the page up to one
3. IF all the pages in the web cache then it is to be marked.
4. If all are marked than unmark all the pages.
5. Discard randomly an unmarked page in the cache.
6. Add time counter with each hit
7. Save the time i.e execution time and repeat step 2.

IV. EXPERIMENTAL RESULTS

A proposed Randomly Page Replacement Caching Algorithm(RPRCA) is developed in windows 7 by using C# .Net (Microsoft visual studio 2010) and the unique identification number is assigned to unique URL‘s to log of proxy server. These numbers is taking as a reference string that suits input to these algorithms. we have also add time counter for counting execution time, and compare hit rate and execution time to all the previous algorithms The results of this algorithm on the basis of hit ratio and execution time are shown in Table 1.

Table 1 Result of number of request and execution time

Number of Requests & Execution time	Number of Hit rate and Execution time Analysis on Different Number of Request is 400			
	Randomly Page Replacement Caching Algorithms			
	FIFO	LFU	LRU	RPRCA
100	3	2	1	5
	683ms	583ms	883ms	329 ms
200	19	18	18	20
	834ms	349ms	529ms	340ms
300	39	35	38	42
	690ms	733ms	816ms	624ms
400	65	65	65	103
	830ms	701ms	680ms	650ms

V. CONCLUSION

This paper is basically focused to explore the increase the hit rate and improve execution time with applying Page replacement algorithm which is best suitable for proxy server. A real trace web reference is fetch with the help of log file details of proxy server. For the simulation numeric reference string was acquired by giving numeric identity to each of the URLs. During simulation we have compared FIFO, LRU and LFU replacement algorithms with this algorithm i.e Randomly Page replacement caching algorithm (RPRCA) . Here we have compared FIFO, LRU and LFU replacement algorithms with this proposed algorithm (RPRCA) We found that RPRCA algorithm 9.35% better than FIFO, 14.35 % better than LFU and 7.6% better than LRU algorithm in terms of **hit rate and execution time**. After comprehensive simulation experiments is summarized that for proxy caching the RPRCA hit ratio and execution time performance better than others existing algorithms.

REFERENCES

1. David A. Malts and Pravin Bhagwat “Improving HTTP caching proxy performance with TCP tap”. Technical report, IBM. March 1998.
2. Abraham Silberschatz And Peter Baer Galvin, “operating system concepts”5th Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA , 1997.
3. John Dilley, Martine Arlitt and Stephane Perret “Enhancement and Validation of Squid’s Cache Replacement Policy” Internet Systems and Applications Laboratory HP Laboratories Palo Alto HPL- 1999-69, May 2009.
4. Ismail Ari, Melani Gottwals Dick Henze ,” SANBOOST: Automate SAN Level Caching in Storage Area Netowkrs “ International Conference on Automatic Computing pp-164-171,2004.
5. Martin Arlitt, Ludmila Cherkasova, John Dilley, Richard Friedrich, Tai J in "Evaluating Content Management Techniques for Web Proxy Caches", published in ACM SIGMETRICS Performance Evaluation Review Volume 27 Issue 4, March 2000.
6. J. Almeida and P. Cao "Measuring Proxy Performance with the Wisconsin Proxy Benchmark", Technical Report, University of Wisconsin-Madison, April 1998.
7. R. Gupta and Sanjiv Tokekar, ”Pair of replacement algorithms MFMR and AF-LRU on L1 and L2 cache for proxy server”. INDICON 2009 publication by IEEE.
8. P. Cao and Irani, “Cost aware WWW Proxy Caching Algorithms” , In roc USENIX Symp. Internet Technology and System Monterey 1997.
9. Yong Zhen Guo, Kotagiri Ramamohanarao and Laurence A. F. Park “Personalized PageRank for Web Page Prediction Based on AccessTime-Length and Frequency” This paper published in 2007 IEEE/WIC/ACM International Conference on Web Intelligence.
10. Gala Golan “Multilevel cache management based on application Hints” computer science department, Technion Haifa 32000, ISRAEL. November 24, 2003.
11. Shiva Shankar Reddy P,Swetha L “Analysis and Design of Enhanced HTTP Proxy Cashing Server“ paper published in International Journal of computer Technology, Volume 2 (3), 537-541.
12. V. Sathiyamoorthi and Dr.Murali Bhaskaran “Data Preprocessing Techniques for Pre-Fetching and Caching of Web Data through Proxy Server” International Journal of Computer Science and Network Security, VOL.11 No.2011.
13. Song Jiang, X. Ding E. Tan and X. Zhang,”DULO: An Effective Buffer Cache Management Scheme to Exploit Both Temporal and Spatial Locality.” Proc. Of the USENIX Symp. File Storage and Technologies 2005.
14. F.J. Gonzalez-Canete, E. Casilari, Ahcia Trivino-Cabrera "Characterizing Document Types to Evaluate Web Cache Replacement Policies," ecumn, pp.3-11, Fourth European Conference on Universal Multiservice Networks (ECUMN'07). 2007 IEEE.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

15. J. L. Hennessy and D. A. Patterson, Computer Architecture: A Quantitative Approach, 3rd Edition, Chapter 5 – Memory Hierarchy Design. Morgan Kaufmann Publishing, 2002.
16. [Online] https://developers.google.com/speed/articles/web-metrics (Accessed 30 June 2015).
17. Y.Zhou and K.Li, "Second Level Buffer cache Management", IEEE Transactions on Parallel and Distributed Systems July2004.
18. L. Bresalu, Cao, S. Shenker , "Web Caching Zipf like Distrubution"1999 IEEE.
19. Vladmir V. Prischepa, "AN Efficient Web Caching Algorithm based on LFU-k replacement policy", Spring Young Researcher's Colloquium on Database and Information System,2004 .