http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 5 (2): 300-304, 2006
ISSN 1812-5638

Efficient Caching Proxy Server Parameters: An Integrated Approach

[1]M.U. Kharat and [2]Narendra S. Chaudhari
[1]Deparment of Computer Sciences and Engineering,
Anuradha Engineering College, Chikhli (MS) India 443201
[2]School of Computer Engineering, Nanyang Technological University, Singapore 639738

Abstract: Caching web proxy server improves the performance of over all web access. Since all the parameters of efficient cache performance are time and locality dependent, no single policy of caching proxy servers is ultimate and applicable for ever. To get more and more efficient performance of caching proxies, history based analysis for a specific locality is found advantageous. Using the history based algorithm popularity of web objects was determined which then further utilized for replacement of cache objects. This approach assures service of updated documents to the clients. Soft caching was used for storing low resolution versions of images which provides efficient and effective services to the clients. Study of various formats of image storing policies have enabled us for getting memory requirements of web objects which turn is closely related to the bandwidth utilization. An integrated approach of studying caching proxy server, recording history of web access, soft caching, improve over all efficiency of web proxy servers.

Key words: Proxy server, web caching, soft caching, bandwidth, cache efficiency

INTRODUCTION

Continuous increase in the number of Internet users has taken exponential shape of growth. It is becoming difficult to adopt this increased number of users and provide satisfactory services to all because of infrastructural limitations on www[1,2]. Caching web object at proxy servers has proven as one of the best alternatives for fast services.

Table 1 shows the exponential growth of internet users for last two decades. The tremendous increase in the internet users is observed in last few years. Bandwidth requirements, placement of web documents, capacity of infrastructure to support bandwidth services, has become major research issues in the last recent years[3].

Caching was initially introduced in the computer systems to provide intermediate memory space between the main memory of the computer systems and the processor. This was with the objective of keeping the programs, data, closer to the processor for faster execution. Furthermore, looking in to the advantages of cache memory, their improvement becomes a major research issue. It is then implemented for internet access by means of caching servers (Web proxy servers). Most of the web servers are placed for internet access with such web proxy servers where the web objects are cached and placed closer to the end users. Cache consistency mechanisms have been included in almost every proxy cache server. Web caching has introduced an effective solution to the problems of traffic congestion, bandwidth requirement and accessing over the web. Efficiency of caching web objects and their access depends on the selected cache updated scheme.

Caching proxy servers improve latency of retrieving web documents and reduce the network traffic on the internet[4,5]. However, with these advantages, the efficiency of cache may reduce over a time. This reduced efficiency depends on several factors such as actual use of caches and accessing cached documents, type of requested document and load on proxy servers that caches the web documents.

Internet access over last twenty five years. Curtsey[3] (Estimated Number of users for 2005 is 500000000).

PROBLEMS WITH CACHING FOR REDUCED EFFICIENCY

The reduction in the efficiency of cache depends on use of cache, access of cached documents, number of users, reference of locality, type of requested document, load on proxy server and its configuration etc.

All these characteristics are subject to change and no fixed policies are available related to above characteristics,

Corresponding Author: M.U. Kharat, Department of Computer Sciences and Engineering,
Anuradha Engineering College, Chikhli (MS) India 443201

Table 1: Exponential growth of internet user

Year	Approximate internet users
1981	300
1985	4000
1990	100000
1995	4000000
1998	10000000
2001	70000000
2002	90000000
2003	200000000
2004	300000000
2005	500000000

Table 2: Estimated proxy server characteristics using history database

Proxy server characteristics parameter	Value
Total number of requests	87458
Number of unique documents	22347
Percentage of unique documents (of total requests)	25.55
Number of one- timers	18511
Percentage of one –timers (of unique documents)	82.83
Total bytes of unique documents(GB)	3
Smallest size of web object(bytes)	32
Biggest size of web object(bytes)	21,809,968
Mean size of web object(bytes)	10,878
Median size of web object(bytes)	3,839

it is unlikely and uncertain that particular configuration and deployment will remain effective over the whole cache life cycle. As well, certain issues must be focused for better and efficient solution of cache efficiency such as:

- Does reference of locality exist with respect to the internet access? What is its impact on internet bandwidth?
- What is the variation in available bandwidth during the course of a cache life cycle? How are number of users accommodated during the course of time?
- What optimum hit rates could be obtained theoretically and practically for better services.
- What performance gains are obtained using currently installed cache policies? How does caching of documents support reference of locality?
- What is the assurity of cached document to be updated one?

HISTORY BASED INTEGRATED APPROACH

Question related to improvement of efficiency are answered with the approach of looking towards cache management strategies. For improvement in efficiency, history based approach of web access is used in our study. We have studied the proxy server setup that is connected and serving to various clients in the network on behalf of desired web server. A log file of every client systems is the target file for study. We obtained several characteristics such as type of web object, its access time total number of bytes required for storing the object, reference count of its access, the client that is requesting for document, total no of hits, total size of web documents that are cached on and served by proxy server (Table 2). Using these parameters[6] from history of log files, we have calculated total number of requests in the specified time, number of unique documents, number of one timers, total bytes received to store various types of documents, smallest and biggest size of documents. In deploying history based approach and to optimize the benefits of web proxy server, we suggest internet managers to reconfigure the existing infrastructure of proxy server for frequent measurements of cache efficiency.

Request made by several users are mainly for image objects of the Internet. Huge amount of memory is required for storing such images on the servers as well as bandwidth for transmission. It is, thus, expected from caching proxies that they should be able to perform recording of images in the cache so that lower versions of images can be stored and made available to the clients. This is termed as soft caching. Storing of lower versions of images is better option than the option that suggests to store higher versions of images or not to store images. This allows users to stop the transfer when sufficient quality image has been downloaded.

To assure updated documents services, it is essential to study the cache replacement algorithms. Depending on the history of web access in the specified environment, web cache replacement algorithm can be implemented. Popular cache replacement algorithms are least recently used, most frequently used least frequently used[7,8]. We suggest study of web proxy servers with the view of web caching, soft caching, image storing strategies and image replacement algorithm.

HISTORY GENERATION ALGORITHM

```
For client internet request
    get request from client;
    {
        record URL, Web pages address;
        record client IP address
        Record, accessdate() and accesstime()
        record size of document()
        set reference count = 1;
        {
```

```
        if document is available in cache
        {
                server it using local cache
                Record sence as hit;
                reference count = reference count + 1
                If document is not available in cache
                {
                        forward request to web server and fetch it in
                        cache if memory space available
                        {
                                if memory space is not available;
                                delete some document form cache
                                 size and store the new document
                        }
                        server the document to the client using its IP
                        address
                }
        }
        With every further request of same document increase
        reference count by 1
    }
}
```

In the other step of efficient cache management we trace for bandwidth availability during the course of a day.

MODELING DECISION FUNCTION FOR SOFT CACHING

History based records give initiations for cache management specifically for documents replacement policies from cache. In defining the scope of proxy server we limit the cache size by allocating the memory space for storing web objects. As the content stored in cache area reaches the cache predetermined limits, cache replacement policies become important for cache management. Obviously using algorithm 1.0 we refer the reference count of web object for its placement in the cache size. Lower is the reference count of web object, higher is the probability of replacement of the web object. Higher reference count of web object indicates popularity of it.

We assume various parameters related to cache management problem as:

With S total cache allotted size and N number of cache content objects;

S(i) : Size of ith object in cache (kb)
t_c : Caching time of object
l_t : Last update time of web object
h(i) : Hits of web object (i)
H(i) : Total no of hits related to all web documents of the cache size
fa(i) : Decision making function for caching
P(i) : Popularity of object (i)

Using conventional method of determining popularity we define it as:

$$\text{Popularity of web objects(i)} = \frac{\text{Hits of web object(i)}}{\text{Total hits of all objects}} \quad (1)$$

$$P(i) = \frac{h(i)}{H(i)}$$

Depending on the percentage metric value of popularity we make decision of whether the object is to be added or replaced. This decision can be modeled as an action function for cache management as described in algorithm 1.

$$f_a = \begin{cases} 1; & \text{if document to be cached} \\ 0; & \text{if document to be removed} \end{cases} \quad (2)$$

CACHE MANAGEMENT POLICY FOR IMAGES

With changing era of Internet access, all the globe is mainly dealing with multimedia documents. Images and audio files on the net require large amount of memory space as compared to text documents. To resolve the problems related to caching of images Eq. 2 is used for making decision of storing the web object in cache. We treat this function faces as hard decision i.e. it either returns value to store or not to store web document. Some

times this decision is disadvantageous even if it is based on history of internet access, as documents are removed from the cache and there is request to such document. In this case complete document is then fetched from web server and cached at proxy for satisfying the clients request.

Instead, soft caching suggests an approach for storing images with their low resolution versions where it consumes limited amount of cache memory. If the purpose of the client is not satisfied with these low resolution images, using reload option of browsers or using video on demand options, client can request for the full resolution of version image. This approach of soft caching enhances the efficiency of caching proxy servers.

Another approach of caching images is based on their formats of storing. We have studied various formats of images and found that for a fixed size of 512X512 resolution, 40% reduction in memory requirements can be achieved using ldf image format and 33.6% of memory size reduction for lwf format. Similar results are obtained for GIF and JPEG image formats[9].

Proxy cache replacement algorithm answer the question (iv) of caching problems for reduced efficiency. Web cache replacement polices based on history reports of internet access leads towards increase in the efficiency of caching. Replacement algorithms that are widely used. We generated using following definitions and parameters.

Latency of web access: It is the time spent between the last byte of request and the first byte of the response. This corresponds to the time a user will have to wait before receiving any data in the browser. It is related to both the network latency and the latency of servers (i.e., web server and proxy server).

Round-trip time: It is the time needed for a request to be completed, from the first byte of the request until the last byte of the response. Value of round trip time leads for either miss or hit to the web object. Negative value of RTT corresponds to miss and positive value corresponds to hit.

Bandwidth savings: These are simply the number of bytes transferred for each transaction including the headers of the request and response. The percentage metric is computed as the ratio between the number of bytes saved (wasted) by a proxy and the total number of bytes for the transaction without the proxy.

Document consistency: The objective of document consistency is at capturing the freshness of the documents returned by the cache. When a cache returns a document it holds, there is no warranty that it has not been updated the original server. The document that is kept in cache must be updated one so that it will improve efficiency of cache by providing updated and correct information

CONCLUSIONS

Looking in to experiential growth of internet users (Table 1), web caching has become a technique for providing efficient services to the users. Web caching at proxy servers level assures about availability of documents at local bandwidth without putting much of burdon on the internet bandwidth and traffic. Analysis at proxy server characteristics from Table 2 shows 71.51% saving of bandwidth if we access 707.756 mb of documents from web servers and proxy server.

Image storing strategies lead towards soft caching approach. It is estimated from the decision function $f_a(i)$ for keeping/removing of document at proxy server. Study of various formats of images enhances the memory requirements of web objects. Jpeg, LDF, LWF are the formats that consumes less memory for images. With fixed resolution of 512X512, average 40% of reduction of memory size for LDF format and 33.6% of memory size reduction for LWF format can be obtained.

An integrated approach with combination of caching of documents at proxy servers, using soft caching based on decision function $f_a(i)$ for documents placement at proxy and image storing formats, we achieve efficient caching performance of the internet with less burdon on bandwidth and internet traffic.

REFERENCES

1. Rainer, H., M.N. Hurber and S. Schroder, 1995. ATM Networks: Concepts, Protocols, Applications. Addison-Wesley.
2. Dhiman, D.C., 2000. High Speed LAN Technology Hand Book. Springer.
3. Douglas, E.C., 2004. Computer Networks and Internets with Internet Applications. 4th Edn., Pearson Education.
4. Daniel, M., 1999. Internet Intranet Engineering: Technologies, Protocols and Applications. Tata Mc-Graw Hill Edition.
5. Anirban, M., D. Eager and C. Williamson, 2000. Temporal locality and its impact on web proxy cache performance. Performance Evaluation, Univ. Sastachewan, C., 42: 187-203.
6. Kharat, M.U. and N.S. Chaudhari, 2005. Parameter selection schemes for web proxy servers and their performance evaluation. Inform. Technol. J., 4: 416-419.

7. Baentsch, M., L. Baum, G. Molter, S. Rothkugel and P. Sturm, 1997. Enhanching the web's Infrastructure. From caching to replication. IEEE Internet Computing University of Kaiscrslaution.
8. Meira, W. Jr., E. Fonseca, C. Murlam and V. Almeida, 1998. Analyzing performance of cache server hierarchies. XVIII International Conference of the Chilean Society of Computer Science, IEEE Catalog No. 98EX212.
9. Chaudhari, N.S. and M.U. Kharat, 2002. Optimal utilisation of bandwidth for internet applications: Proceedings of International Conference on e-manufacturing, MANIT Bhopal 17-19 Nov., pp: 109-113.