Proxy Caching Impact in Performance Improvement of Server and Web Access

Amna Hashim Mohammed Ali1, Dr.AminBabikerAbdalnabi Mustafa2, Dr.AshrafGasim AL Seed Abdallah3

Department of Communication, Faculty of Engineering / Al Neelain University, Sudan1
Department of Communication, Faculty of Engineering / Al Neelain University, Sudan2
Department of Communication, Faculty of Engineering / Future University, Sudan3

ABSTRACT: Proxy server has many benefits; one of them is proxy cache. Cache proxy is a function of a proxy server, which caches retrieved Web pages on the server's hard disk; so it achieves resources available for users in the next time that page is requested. Cache proxy enhance QoS through achieving shorter response time, reduced bandwidth requirement, reduced load on servers, reduced access control complication and logging. This paper focuses on impact of reducing response time and reducing load on servers performance.

KEYWORDS: QoS, HR, HTTP.

I. INTRODUCTION

Caching is an effective technique used in computer system for decades to enhance performance and achieving QoS. Proxy caching differs substantially from the traditional cache used in processors. World Wide Web applications have two type of caches that differ from the traditional caching called non-uniformity of the object sizes and non-uniformity of the cost of cache misses, as opposed to the traditional caching where all cache blocks have the same size and require the same amount of time to be retrieved in case of cache misses.

The traditional metrics for measuring caching efficiency has been hit-ratio (HR). HR is defined in mathematical form, as the number of requests satisfied by the cache divided by the total number of requests. Obviously, HR is not an appropriate metric to measure performance of proxy caches, because of the non-uniformity of object sizes and non-uniformity cost of misses. Byte Hit Ratio (BHR) is defined as the number of bytes found in the cache divided by the total number of bytes requested within the observation period. The BHR takes into consideration the non-uniformity of the object sizes, but it fails to consider the non-uniform cost of misses. [1][2]

One of the effective parameter in proxy server performance is Page response time, which defined as the time between a web browser requests from your server to the time when web server responds to that request. That mentioned scenario may be repeated many times per second, so we can defund it as request/sec. Also the average response time is amount of time it takes response to be complete.

“WWW proxy caching attempts to improve performance in three ways. First, caching attempts to reduce the user-perceived latency associated with obtaining Web documents. Latency can be reduced because the proxy cache is typically closer to the client than the content provider is. Second, caching attempts to lower the network traffic from the Web servers. Network load can be lowered because documents that are served from the cache typically traverse less of the network than when they are served by the content provider. Finally, proxy caching can reduce the service demands on content providers since cache hits need not involve the content provider.”[4]

Cache technique has three different kinds in the context of websites called Private Cache, Forward Proxy Servers and caching proxy server. In Private Cache If a browser has already sought a web resource and needs the same resource again, it knows (through HTTP Headers) to find it through a local copy instead of having to request one remotely. In this case, the browser will receive a 304 message from the server (which is the signal for the browser to use its private cached). Ultimately, this kind of cache has nothing to do with Enfold or IIS. By default, the http headers for Enfold
Proxy give Age=0, and that setting will always cause EP to check for a cached version. The Forward Proxy Servers keeps local copies of frequently requested resources, allowing large organizations and ISPs to significantly reduce their upstream bandwidth usage and cost, while significantly increasing.

II. METHODOLOGY

In this paper, we simulated the network traffic and Web Caching and Data Compression to investigate Improvement of Web Access and Server Performance by Observing the Responses time, Hit-rate and Request/sec. we use OPNET v.14 to build three scenarios of a network with main clients accessing a Web server.

III. SYSTEM DESCRIPTION

The network consists from main office network, which includes of web server (The main office) and sub network as depicted in Fig (1). Other client network (Portsudan office) includes LAN network .Main and client networks connected through router as shown in Fig (2). The client office consists of LAN, Server and router as show in fig (3).
Fig (3): Client Office network

Simulation implemented depending on two scenarios, which represent all expected situations for the system. The scenarios are as below:

Scenario 1: The NoCache_Comp Scenario, The clients network created (Portsudan Office) access the Web server in the Main office. Since the link between the two subnets is slow and to improve the response time of accessing the Web pages, we will create another scenario for the same network where we will utilize the feature of compressing the payload of the IP datagrams. In OPNET, compressing the payload is referred to as Per-Virtual Circuit Compression.

Scenario 2: The Cache _No Comp Scenario: Another option to improve the Web page access response time is to utilize a cache server. Layer 4 Switch is a switch capable of redirecting application traffic based on the application protocol. Normally this switch used in conjunction with HTTP traffic to simulate transparent Web caching scenarios as shown in fig (4).

Fig (4): The Cache_NoComp Scenario
After simulation implemented the obtained results, the graphs show the comparison between three building scenarios and explain the improvement points.

**Graph (1)**

Graph (1) shows the system running without choosing any option of analysis, the graph explains that the system in the beginning seemed to be stable and changes appeared depending on the request rate.

**Graph (2)**

Graph (2) has shown the relation between hit rate in cache server and client HTTP page response time, the graph explains the instability of Hit – rate in the beginning. After a while Hit – rate became stable.
Graph (3) shows the result in time average refinances, we see in (No Cache) the load Request time in server higher than the load Request time in (cache); so the caching improve performance in server.

Graph (4) shown the same observe with sample sum; to be more clear.

Graph (5)
As graph (5), the page request increase in server but the response time and hit rate in proxy stable. That mean the client did not affect with the high load of requests in server, because some page which request came from proxy without going back to the original server and connect the client with it, the client cannot know this page came from the original server or not; so the final result is significant improvement in performance.

V. CONCLUSION

cache proxy technique has significant contribution in improving Web Access and Server Performance by reducing Page response time, the result represent in decreasing delay, which will impact directly in performance improvement through enhancing bandwidth utilization, reducing load on servers and reducing access control complication.

REFERENCES