10 th International Conference on Innovation and Research in Engineering Sciences
GEORGIAN INTERNATIONAL ACADEMY OF SCIENCES AND STUDIES
11 November 2021 -TBILISI GEORGIA
www.icires.ir
info@icires.ir

Reduce response time in content delivery networks with a hybrid method

Mohammad Malek barmi
Department of computer science
Islamic Azad University of Damghan
Damghan, Iran
Mohammad.Malekbarmi@gmail.com

***Abstract*—One of the effective factors in the efficiency of content distribution networks is finding an optimal placement and proper content delivery. Most previous research has been done separately on these two methods and in general, caching in proxy and duplicate server systems have been studied as one of the mechanisms in content delivery networks. In this paper, with combining replication and caching techniques a hybrid method is presented. This method uses replication and caching benefits and in each iteration, all possible contents are evaluated and according to the conditions, the best ones are selected and replicated. The proposed method is tested in a simulation environment and the results show that compared to the methods based on replication and caching, it has been able to reduce the response time by an average of 20%.**

I. INTRODUCTION

With the spread of the Internet and the increasing need of users for services under it, popular websites suffered heavy traffic, so the quality of services provided declined [1]. Content delivery networks were proposed to address these issues. These networks replicate and cache content from the main server to surrogate servers scattered in different geographical locations [2]. In this way, the requested contents of the user are delivered from the server of the representative close to him. It is clear that when node-based systems reduce network communication traffic, system reliability, efficiency, and scalability increase. At the same time, content delivery networks are one of the best ways to prevent bottlenecks, increase response time or fail to provide services to the user by using replicate content and caching techniques [3]. In this case, static data is often replicated in surrogate servers and it can be reducing the traffic load on the main server and the contents are delivered faster and more securely to the user. In addition, surrogate and caching techniques cause requested dynamic contents to be stored on the surrogate replacement servers in order to respond more quickly to future requests [4].

Accordingly, in this paper, by combining caching and replication techniques, a greedy method for optimal content placement on surrogate CDN servers is presented. The structure of the article is as follows: in the second part, the background of the work done in the field of content delivery in CDN is reviewed, then in the third part, the proposed method is presented and in the final part, the results of the evaluation are described.

II. LITERATURE REVIEW

Considering the importance of content placement in CDN, various studies have been done in this context. In [5], a greedy algorithm based on Greedy has been presented to improve content placement by determining the best content at each replica. The tests show that the simple LRU can improve the response time for HTTP requests efficiently. In [6], in addition to emphasizing the importance of placement algorithms in increasing efficiency of the content delivery networks, an optimal model with a server storage has been presented for replica placement. The main part of this model is designed based on the minimum cost model for the network traffic, and its main objective is to reduce the total network cost. Finally, the algorithm is examined using a simulator and its cache hit rate is given. In [7], the caching methods are combined to manage the storage space of the servers optimally. In this paper, a nonlinear classifier based on learning has been presented that classifies the objects of the server storage disk in specific intervals based on the scores obtained by the mentioned function. Simulations showed that this method has