



Various Techniques of Web Pre-Fetching

Rupinder Kaur, Vidhu Kiran

Research Scholar, Assistant Professor

Department of Computer Science and Engineering

Jan Nayak Ch. Davilal College of Engineering, Sirsa, India

Abstract— *The explosive growth of the Internet and the World Wide Web results in network congestion and server overloading. Web traffic reduction techniques are necessary for accessing the web sites efficiently with the facility of existing network. Web pre-fetching techniques and web caching reduces the web latency. In web pre-fetching web pages and web objects are pre-fetched into the proxy server cache. Web caching is used to reduce the network traffic and server overload. This paper describes the various schemes of web pre-fetching and method to improve caching.*

Keywords— *Web Caching, Web Pre-Fetching*

I. INTRODUCTION

In order to reduce latency pre-fetching is required necessarily. The term pre-fetching means fetching web pages in advance by proxy server. Web pre-fetching exploits the spatial locality of web pages, those pages that are linked with current page will be accessed with higher probability than other pages. Web pre-fetching can be applied in a web environment as between clients and proxy server [3], due to this proxy starts feeds pre-fetched web objects from its cache to the clients so that there won't be extra internet traffic. This paper describes about the various pre-fetching techniques, and also about predictability of web objects to be pre-fetched. Pre-fetching is used as an attempt to place data close to the processor before it is required, in order to eliminating cache misses as possible. Pre-fetching is the means to anticipate probable future requests and to fetch the most probable documents, before they are actually requested. It is the speculative retrieval of a resource into a cache in the anticipation that it can be served from the cache in the near future, thereby decreases the load time of the object. This paper describes about the various pre-fetching techniques, how they predict the web object to be pre-fetched and what are the issues involved in these techniques.

II. CLASSIFICATION OF WEB PRE-FETCHING TECHNIQUES

Categorization of web pre-fetching can be done on the basis of its long term and short term benefits. These are classified according to their benefits as:

A. Short Term Pre-Fetching Policy

Pre-fetching used in traditional short term approach was based on accessing history to predict and Pre-fetch objects likely to be referenced in the near future. In short term policies, based on the client's recent access history, objects that are likely to be referenced in the near future are pre-fetched. Future requests are predicted to the cache's recent access history. Based on these predictions, clusters of Web objects are pre-fetched. Several short-term pre-fetching policies exists nowadays namely Predictive Web pre-fetching, semantic web pre-fetching and proxy web pre-fetching. These pre-fetching methods are based on Markov models, which are used for modeling and predicting user's browsing behavior over the Web.

B. Long Term Pre-Fetching Policy

The long-term pre-fetching uses long-term steady-state object access rates and update frequencies to identify objects to replicate to content distribution locations. In long term pre-fetching, caching is done on comparing the demands by clients, it increases network bandwidth and disk space costs but may benefit a system by improving hit rates. In this policy, objects are updated and pre-fetched based on long-term global access and update patterns. Global object access pattern statistics (e.g., objects' popularity, objects' consistency) are used to identify valuable objects for pre-fetching. In this scheme a specific pattern is followed, the objects with higher access frequencies and no longer update time intervals are more likely to be pre-fetched. There are several types of long term pre-fetching policies exists namely frequency based pre-fetching, greedy dual-size pre-fetching and popularity based pre-fetching.

III. WEB PRE-FETCHING TECHNIQUES

Web pre-fetching is the process of deducing client's future requests for web documents and getting that document into the cache. The main advantage of employing pre-fetching is that it prevents band width underutilization and hides part of latency. The Web pre-fetching can be classified as described below:

A. Domain Top Approach

Domain top approach is proposed by Seung Won Shin et al. for web pre-fetching, this approach combines the proxy's active knowledge of most popular domains and documents[3]. In this approach proxy is responsible for calculating the most popular domains and most popular documents in those domains, then prepares a rank list for pre-fetching. Main aim of this method is to increase the hit ratio by the proxy pre-fetching and to put a little burden on the proxy and the network. In our scheme, proxy finds the popular domains using access profiles and searches the popular documents in each domain. Based on these Top-Domain and Top-Documents, proxy makes the Rank list for pre-fetching. In addition, this approach can be implemented without changes of server and client.

B. Link Pre-fetching

Fisher et al. proposed a server driven approach for link pre-fetching [15]. Link pre-fetching is a proprietary syntax to give web browser a hint about documents that it should pre-fetch because that can be demanded by the client again. In this approach browser follows special directives from the web server or proxy server that instructs it to pre-fetch specific documents. This mechanism allows servers to control the contents to be pre-fetched by the browser. It utilizes browser's idle time to download documents that the user might visit in the future. Set of pre-fetching hints are provided by the web page to browser and after completion of loading the page, browser starts pre-fetching specified documents and stores them in its cache. When the user visits one of the pre-fetched documents, it can be served up quickly out of the browser's cache. When the browser is idle, it observes these hints and queues up each unique request to be pre-fetched.

C. Top 10 Approach

Evangelos P. Markatos et al. proposes a top 10 approach to pre-fetching on the web, in which the server calculates the list of most popular documents [2]. This approach is easy to implement in a client server architecture. It considers the frequency of access for predicting the web object, not the client characteristics on the web. This approach combines the servers' active knowledge of their most popular documents (their Top-10) with client access profiles. Based on these profiles and clients requests, servers forward them, regularly, their most popular documents. The Top-10 approach to pre-fetching is based on the cooperation of clients and servers to make successful pre-fetch operations. The server side is responsible for periodically calculating a list with its most popular documents (the Top-10) and serving it to its clients.

D. Dynamic Web Pre-fetching

In dynamic web pre-fetching technique [5], a preference list is stored in proxy server's database, which is known as a user's preference list. That means each user can keep a list of sites to access immediately. Intelligent agents are used for parsing the web page, monitoring the bandwidth usage and maintaining hash table, preference list and cache consistency. It controls the web traffic by reducing pre-fetching at heavy traffic and increasing pre-fetching at light traffic. Thus it reduces the idle time of the existing network and makes the traffic almost constant. A hash table is maintained for storing the list of accessed URLs and its weight information. Depending upon the bandwidth usage and weights in the hash table, the prediction engine decides the number of URLs to be pre-fetched and gives the list to pre-fetch engine for pre-fetching the predicted web pages. After pre-fetching, the proxy server keeps the pre-fetched web pages in a separate area called pre-fetch area. The major advantage of using this technique is that, number of links to be pre-fetched is based on the current band width usage. This technique increases overhead to proxy or client.

E. Model based Predictive Pre-fetching

This pre-fetching approach is proposed by Yang et al. This is a model based approach in which integration of web-caching and web pre-fetching model is used [22]. The prediction model used in this is based on the statistical correlation between web objects. The prediction model is time based, prediction window represents some specific time period than number. The algorithm constructs a logical graph called correlation graph, which shows the correlation between web objects and pre-fetch web objects that are highly correlated to a currently requested object. They developed an integrated caching and pre-fetching algorithm, Pre-GDF. This algorithm is based on the algorithms GD-Size [23] and its enhancement GDSF [24]. The key components in the algorithm are replacement manager, pre-fetching agent, prediction queue and cache.

F. Interactive Pre-fetching Scheme

All the hyper links and inline images in linked pages are fetched in this scheme. As it retrieves all the hyper links, there are only 20 % chances of miss, hit rate of 80% is possible. The disadvantage is it increases the load on to the host, and requires a lot of memory to store the pre-fetched web pages.

G. A Keyword based semantic pre-fetching approach

This approach finds out semantic preferences by analyzing keywords in URL anchor text of previously accessed documents in different categories [4][17]. It proposes a key word based semantic pre-fetching, in which prediction of future requests are based on semantic preferences of past retrieved web documents. It assumes a proxy is running behind the web browser keeps track of client's characteristics and find out semantic relation between the documents [4]. This technique uses a neural network based semantic model, which has the capability of self learning. So this technique is capable of pre-fetching URLs that have never been accessed.

H. Adaptive pre-fetching Scheme

Chen and et. al[27] proposed an adaptive pre-fetch scheme, in which dynamically adjust the pre-fetch aggressiveness in web servers and uses a threshold to adjust the aggressiveness of pre-fetching. Fagni and et. Al [28] proposed uses an approach for boosting the performance of search engine by exploiting the spatial and temporal locality present in the stream of processed queries. Jiang and et al. proposed an adaptive pre-fetch scheme, in which the number of files to be pre-fetched depends on user access history and network conditions[26]. Adaptive pre-fetch scheme are developed to adapt user's browsing history and habits [25]. This scheme consists of two modules: prediction module and threshold module. The prediction module updates the history and also computes the access probability of each file. Files whose access probabilities greater than or equal to the pre-fetch threshold are only pre-fetched.

I. Markov model for predicting web access

Markov model and hidden Markov model are used for predicting the web pages to be pre-fetched based on the web access probability. Traditional Markov model takes a sequence of web pages accessed by the user as an input and predict the next page to be accessed by the user. Low order Markov models can't accurately predict the subsequent requests of user. If higher order Markov models are used the memory required increased rapidly. To solve this, Xing Dongshan et. al proposed hybrid order tree like Markov model, which can predict web access precisely, providing high coverage and good scalability[12]. Xin Jin et. al[13] proposed a novel web pre-fetching approach based on Hidden Markov Model. The HMM is used for analyzing user's browsing history and make semantic based web pre-fetching decisions. These models are based on historical data, so these are unable to predict a new web request.

J. Data Pre-fetching

Data pre-fetching is a data access latency hiding technique, which decouples and overlaps data transfers and computation. In order to reduce CPU stalling on a cache miss, data pre-fetching predicts future data accesses, initiates a data fetch, and brings the data closer to the computing processor before it is requested. A data pre-fetching strategy has to consider various issues in order to mask data access latency efficiently. It should be able to predict future accesses accurately and to move the predicted data from its source to destination in time.

K. Content Pre-fetching

The content sensitive data pre-fetching is a new hardware technique for data pre-fetching. This technique is simple, effective, and realizable content sensitive data pre-fetcher can be built as an on-chip component of the memory system, and will scale well with future processor designs. One of the major findings is that the content pre-fetcher enhanced memory system is capable of delivering timely pre-fetches to fully suppress demand load misses for nearly all applications. The large percentage of full content pre-fetch hits fully validates the decision to place the content pre-fetcher on chip and positively indicates that the proposed memory system that utilizes prioritized memory traffic and pre-fetch reinforcement is well balanced.

L. Context based Pre-fetching

The Context based Pre-fetching technique is used for predicting useful pre-fetches when a navigational object-oriented interface was implemented on a relational DBMS. When implementing persistent objects on a relational database, a major performance issue is pre-fetching data to minimize the number of round-trips to the database. This paper proposes the use of the context in which an object is loaded as a predictor of future accesses, where a context can be a stored collection of relationships, a query result, or a complex object.

IV. CONCLUSION

Web pre-fetching acts complementary to caching, it can significantly improve cache performance and reduce the user-perceived latency. Lot of research is going on in web pre-fetching in various directions. In this paper various web pre-fetching techniques and other directions of web pre-fetching are analyzed and discussed. The web pre-fetching scheme focus on the property spatial locality of web objects. These techniques are applied to reduce the network traffic and improve the user satisfaction. It is applied in wireless networks and efficiently improves the data management.

REFERENCES

- [1] http://shika.aistnara.ac.jp/products/wcol/tech/p_concept.html dated 7/9/2011.
- [2] Evangelos P. Markatos & Catherine E. Chronaki, (1998) "A Top-10 Approach to Pre-fetching on the Web, Proceedings of INET' 98 (The Internet Summit).
- [3] Seung Won Shin, Byeong Hag Seong & Daeyeon Park, (2000) "Improving World-Wide- Web Performance Using Domain-Top Approach to Pre-fetching", Fourth International Conference on High- Performance Computing in the Asia-Pacific Region vol. 2, pp. 738-746.
- [4] Cheng-Zhong Xu & Tamer I. Ibrahim, (2004) "A Keyword-Based Semantic Pre-fetching Approach in Internet News Services", in IEEE Transactions on Knowledge and data engineering, vol.16, pp. 601-611.
- [5] Achuthsankar S. Nair & J. S. Jayasudha, (2007) "Dynamic Web Pre-fetching Technique for Latency Reduction", International Conference on Computational Intelligence and Multimedia Applications.
- [6] Reinhard P. Klemm, (1999) "WebCompanion: A Friendly Client-Side Web Pre-fetching Agent", IEEE Transactions on knowledge and data engineering, pp. 577-594.

- [7] Chii Chang, Sea Ling & Shonali Krishnaswamy,(2011) “Proactive Mobile Web Service Provision Using Context-Awareness”, in 8th IEEE International Workshop on Managing Communications and Services, pp. 69-74.
- [8] Zhimei Jiang & Leonard Kleinrock,(1998) “Web Prefetching in a Mobile Environment”, in IEEE Personal Communications, pp. 25-34.
- [9] Sarina Sulaiman, Siti Mariyam Shamsuddin & Ajith Abraham, (2009)“Rough Neuro-PSO WebCaching and XML Prefetching for Accessing Facebook from Mobile Environment”, in World Congress on Nature & Biologically Inspired Computing.
- [10] Sarina Sulaiman, Siti Mariyam Shamsuddin & Ajith Abraham,(2010) “Intelligent Mobile Web Prefetching (IMWeP) Using XML Technology”, 6th IEEE International Conference on Next Generation Web Services Practices, pp. 475-480.
- [11] Sarina Sulaiman, Siti Mariyam Shamsuddin, Ajith Abraham, (2009)“Rough Web Caching”, Springer Verlag, pp. 187-211.
- [12] Xing Dongshan & Shen Junyi,(2002)”A new Markov model for web access prediction”, pp. 34-39.
- [13] Xin Jin & Huanqing Xu, (2003)“An Approach to Intelligent Web Pre-fetching Based on Hidden Markov Model”, in IEEE Conference on Decision and Control , pp. 2954-2958.
- [14] Weifeng Zhang, Baowen Xu, William Song, Hongji Yang & Kecheng Liu,(2000) “Data Mining Algorithms for Web Pre-Fetching”, Proceedings of IEEE International Conference on Web Information Systems Engineering.
- [15] Darin Fisher & Gagan Saxena, (2003)“Link Pre-fetching in Mozilla : A Server-driven Approach”, Proceedings of 8th International Workshop on Web Content Caching and Distribution.
- [16] <http://www.w3.org/protocols/rfc2616/rfc2616.html> dated 01/02/2011.
- [17] Cheng-Zhong Xu &Tamer I. Ibrahim,(2003) ”Towards Semantics-Based Pre-fetching to Reduce Web access latency ”, Proceedings of the Symposium on Applications and the Internet.
- [18] Themistoklis Palpanas & Alberto Mendelzon,(1998) “Web Pre-fetching Using Partial Match Prediction”, Proceedings of 4th Web Caching Workshop.
- [19] Xin Chen & Xiaodong, Zhang,(2003) “A Popularity-Based Prediction Model for Web Pre-fetching”,IEEE Computer, Vol. 36, No. 3, pp. 63-70.
- [20] Venkata N. Padmanabhan & Jeffrey C. Mogul,(1993) “Using Predictive Pre-fetching to Improve World Wide Web Latency”, Computer Communication Review journal, Vol. 26, No. 3, pp. 22-36.
- [21] Yingyin Jiang, Min-You Wu & Wei Shu,(2002) “Web Pre-fetching: Costs, Benefits and Performance”, Proceedings of 7th International Workshop on Web Caching & Distribution.
- [22] Qiang Yang & Zhen Zhang,(2001) “Model Based Predictive Pre-fetching”, Proceedings of IEEE International Workshop on Database and Expert Systems Applications.
- [23] P.Cao & S. Irani ,(1997) “Cost-aware WWW proxy and caching algorithms ”,in USENIX Symposium on Internet Technologies and Systems.
- [24] L.Cherkasova,(1998) “ Improving WWW proxies performance with Greedy-Dual-Size-Frequency Caching Policy”, in HP Technical report.
- [25] Brian D. Davison,(1999) “Adaptive Web Pre-fetching”, Proceedings of 2nd Workshop on Adaptive Systems and User Modeling on the World Wide Web, Toronto,pp. 105-106.
- [26] Zhimei Jiang & Leonard Kleinrock,(1998) “An Adaptive Network Pre-fetch Scheme”, IEEE Journal on Selected Areas in Communications, Vol. 16, No. 3, pp. 358-368.
- [27] Xin Chen, Xiaodong Zhang,(2003) “Accurately Modeling Workload Interactions for Deploying Pre-fetching in Web Servers”, Proceedings of IEEE International Conference on Parallel Processing.
- [28] Tiziano Fagni, Raffaele Perego, Fabrizio Silvestri, Salvatore Orlando, “Boosting the Performance of Web Search Engines: Caching and Pre-fetching Query Results by Exploiting Historical Usage Data”, ACM Transactions on Information Systems, Vol. 24,No. 1, pp. 51-78.
- [29] Ravi Kokku, Praveen Yalagandula & Arun Venkataramani, Mike Dahlin,(2003) “NPS : A Non- Interfering Deployable Web Pre-fetching System”, Proceedings of USENIX Symposium on Internet Technologies and Systems.
- [30] Qiang Yang, Haining Henry Zhang & Tianyi Li,(2001) “Mining Web Logs for Prediction Models in WWW Caching and Pre-fetching”, Proceedings of 7th ACM International Conference on Knowledge Discovery and Data Mining, pp. 473-478.
- [31] Yin-Fu Huang & Jhao-Min Hsu,(2005) “Mining Web Logs to Improve Hit Ratios of Pre-fetching and Caching”, Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence.
- [32] Bin Lan, Stephane Bressan, Beng Chin Ooi & Kian-Lee Tan, (2000) “Rule-Assisted Pre-fetching in Web-Server Caching”, Proceedings of ACM Conference on Information and Knowledge Management, pp. 504-509.
- [33] Alexandros Nanopoulos, Dimitrios Katsaros & Yannis Manolopoulos,(2003) “A Data Mining Algorithm for Generalized Web Pre-fetching”, IEEE Transactions on Knowledge and Data Engineering, Vol. 15, No. 5, pp. 1155-1168.
- [34] Areerat Songwatta,(2008) “Mining Web Logs for Prediction in Pre-fetching and Caching”, Proceedings of IEEE International Conference on Convergence and Hybrid Information Technology, pp. 1006-1011.