

Fast Speculative Search Engine on the Highly Parallel Computer EM-X

Hayato YAMANA
Computer Science Div.
Electrotechnical Laboratory
www.etl.go.jp/~yamana/

Hanpei KOIKE
Computer Science Div.
Electrotechnical Laboratory
www.etl.go.jp/~hkoike/

Yuetsu KODAMA
Computer Science Div.
Electrotechnical Laboratory
www.etl.go.jp/~kodama/

Hirofumi SAKANE
Computer Science Div.
Electrotechnical Laboratory
www.etl.go.jp/~hsakane/

Yoshinori YAMAGUCHI
Computer Science Div.
Electrotechnical Laboratory
www.etl.go.jp/~yamaguti/

1 Abstract

This paper presents the new World Wide Web search engine called "Fast Speculative Search Engine" that uses speculative execution on multiprocessor systems to shorten the total time to retrieve information from the WWW.

The proposed search engine predicts the user's next queries and initiates the searches with the predicted queries before receiving them to accelerate narrowing the search space. This kind of speculation is classified as the data value speculation [2], which are mainly studied as the scheme to extract the instruction level parallelism in a processor. However, there have been no systems that adopt such speculation on multiprocessor systems.

We have implemented the fast speculative search engine using the data speculation on the EM-X[4] which is shown in Fig.1(1). The EM-X, which consists of 80 processors, is a highly parallel computer which can tolerate communication latency by using low latency communication and multithreading. The peak performance of the EM-X is 1.6 GIPS / 3.2GFLOPS and the point to point network throughput is 37.2 MB/s. On the EM-X,

time when there are many search requests, and provide maximum search service at free time when there are small number of search requests. We call such a controlling scheme as Unlimited Speculative Execution[1].

The search mechanism is shown in Fig.1(2). When the system receives the 1st query, the system initiates the search normally. At the same time, the system calculates "Topic Index Words (TIWs)" which are the words included in the retrieved search space of the 1st query. The TIWs are selected carefully as each of TIWs is included in less than 80% of the search space. After that, the system returns the search results and TIWs.

As soon as returning them, the system speculates the next searches using the 1st query and the TIWs when there exit idle processors. Since over 80% of the query transformations are adding/deleting words, changing to synonym or using same words[3], we use these tendencies to initiate the variety of speculative searches.

In the current implementation, the speculative searches continue until the system receives the users' next query or the speculation reaches the 2nd stage - the speculative search for the query after the next query.

The experimental result, using the data set of the web documents in ETL, shows that the 42% of user's queries hit on the speculative searched results.

We will show the interactive WWW information retrieval using WWW browsers connected to the EM-X located at ETL in Japan by the Internet. We also show the speculation process to accelerate narrowing the search space.

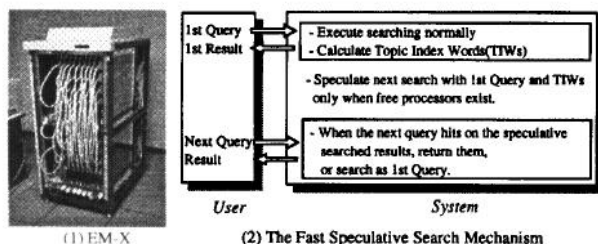


Figure 1: The EM-X and Search Mechanism

idling processors are used to predict the next queries and no predictions are made when all processors are busy. Thus, we can provide minimum search service at busy

Permission to make digital/hard copy of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copying is by permission of ACM, Inc. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or fee. SIGIR'98, Melbourne, Australia © 1998 ACM 1-58113-015-5 8/98 \$5.00.

References

- [1] H.Yamana, M.Sato, Y.Kodama, H.Sakane, S.Sakai, and Y.Yamaguchi: *A Macrotask-level Unlimited Speculative Execution on Multiprocessors*, Proc. of ACM Int. Conf. on Supercomputing'95, Barcelona, Spain, pp.328-337, 1995.
- [2] M.H.Lipasti and J.P.Shen: *Exceeding the Dataflow Limit via Value Prediction*, IEEE Micro, Vol.29, pp.226-237, 1996.
- [3] P.D.Bruzs and S.Dennis: *Query Reformulation on the Internet: Empirical Data and the Hyperindex Search Engine*, Proc. of RIAO'97, Montreal, pp.408-499, 1997.
- [4] Y.Kodama, H.Sakane, M.Sato, H.Yamana, S.Sakai and Y.Yamaguchi: *The EM-X Parallel Computer: Architecture and Basic Performance*, Proc. of Int. Symp. on Computer Architecture '95, pp.14-23, 1995.