Expanding the Scope of Prefetching through Inter-Application Cooperation

Amy S. Hughes and Joseph D. Touch
USC / Information Sciences Institute
4676 Admiralty Way, Suite 1001
{ahughes, touch}@isi.edu

Prefetching web documents is a way to reduce user-perceived latency. Web prefetching works best when web caches (browsers and proxies) have information to predict which documents should be preloaded into the cache. Inter-application cooperation can increase the amount of predictive information available to a cache. Web context can drive web prefetching, or can be used to drive prefetching in related caches for different semantic domains, *e.g.*, the DNS [1] [2]. Although this is similar to hinting, where one application pushes information to another, other means of sharing are possible. Such inter-application cooperation expands DNS prefetching beyond the use of DNS hints alone, providing entirely new predictive information. In turn, inter-application cooperation, can help increase the amount of predictive information available to web caches. This work presents an example, under development, of e-mail-driven web prefetching.

Inter-application cooperation means cooperation between different applications in different semantic domains, *e.g.*, web and e-mail, rather than between different application programs for the same semantic domain. Most examples of cache cooperation involve caches of the same type, *e.g.*, a web proxy cache hierarchy. *Intra-application cooperation* creates a larger, distributed cache in the same semantic domain. In contrast, caches that belong to different applications (*e.g.*, web caches and DNS caches) can cooperate with benefits to one or both caches. When the two caches cooperate, new information is available to drive prefetching.

Prefetching requires some form of predictive information. This information can be derived from user behavior, though web documents are special because inline links make them self-referential. These links represent requests that the user might make in the future. Even when a specific web page cannot be prefetched because it cannot be cached, other related items, such as DNS records, can still be prefetched. This is useful because there are a large number of web documents for which prefetching is not productive, such as dynamic pages. While DNS prefetching relies heavily on web context, it is not web prefetching. Instead, information from web documents is used to predict future use and trigger prefetching across semantic domains. The web cache provides predictive information that was previously unavailable to the DNS. Cooperation with the web cache widens the scope of the DNS cache and enables prefetching as a new optimization.

Like a DNS cache, a web cache can find new predictive information by including new information sources in its search. To do so, it must consider all sources of URLs that are available to the user. They fall into three categories: URLs directly available to the browser, URLs available elsewhere in the host, and URLs received completely out of band (out-of-the-blue guess, watching TV, reading a newspaper, seeing a billboard). Some of these sources are inaccessible to the cache and each source has a different probability of yielding documents that might be good candidates for prefetching.

Of the URLs available in a host, many are found in documents such as e-mail messages and word processing documents. E-mail messages contain URLs that users are likely to request. By their nature, these URLs should be cacheable, because they generally point to static documents. The time between downloading an e-mail message and accessing a URL within it should provide ample time to prefetch the document. In an effort to assess the utility of such cooperation, we have instrumented the Mozilla browser to collect trace logs of URLs within e-mail messages, and to track how the user interacts with these URLs. The information provided by the traces will allow us to characterize the benefits of cooperation. Eventually we will extend the browser to prefetch these URLs to evaluate the effectiveness of intra-application cooperation.

This work examines inter-application cooperation and its applications to web caching. Web-DNS cooperation is an example of this idea in use. E-mail-web cooperation is under development. Inter-application cooperation may extend to new situations as the use of caches to improve user-perceived latency increases.

An extended version of this paper is available as a technical report. For a copy, please contact the authors. The modified Mozilla software, which creates the traces to analyze, can be obtained at http://www.isi.edu/~ahughes/cachecooperation.html.

REFERENCES

[1] E. Cohen and H. Kaplan. “Proactive caching of DNS records: Addressing a performance bottleneck”, in Proceedings of the Symposium on Applications and the Internet, January 2001, to appear.

[2] A. S. Hughes and J. Touch. “Cross-Domain Cache Cooperation for Small Clients”, in Proceedings of NetStore ’99, Seattle, Washington, October 1999.