

ABSTRACT

The sharing of caches among Web proxies is an important technique to reduce Web traffic and alleviate network bottlenecks. Nevertheless it is not widely deployed due to the overhead of existing protocols.

As the tremendous growth of the World Wide Web continues to strain the Internet, caching has been recognized as one of the most important techniques to reduce bandwidth consumption. In particular, caching within Web proxies has been shown to be very effective. To gain the full benefits of caching, proxy caches behind a common bottleneck link should cooperate and serve each other's misses, thus further reducing the traffic through the bottleneck. That process was called "Web cache sharing."

A new cache sharing protocol called "Summary cache." Under this protocol, each proxy keeps a compact summary of the cache directory of every other proxy. When a cache miss occurs, a proxy first probes all the summaries to see if the request might be a cache hit in other proxies, and sends a query messages only to those proxies whose summaries show promising results. The summaries do not need to be accurate at all times. If a request is not a cache hit when the summary indicates so (a false hit), the penalty is a wasted query message. If the request is a cache hit when the summary indicates otherwise (a false miss), the penalty is higher miss ratio.

Two factors contribute to this protocol's low overhead: the summaries are updated only periodically, and the directory representations are very economical, as low as 8 bits per entry.