Design & Implementation of Methods Enchancing the Performance of WWW Proxies

Michail D. Flouris

Master of Science Thesis

Computer Science Department
University of Crete

Abstract

The quality of service (QoS) problem of the World Wide Web and the protection of intranets behind firewalls, lead to the increasing use of WWW Proxies (or WWW Caches) as the only long-term solution. However, due to the exponential growth of the Web, there is an increasing demand for higher performance WWW Proxies with larger WWW object caches, not only in main (dynamic) memory but also on magnetic disks. Contrary to this demand, several research studies have proved, not only that modern WWW proxy systems suffer severe performance problems, but also that the main cause of these performance problems is the magnetic disk I/O induced by the WWW object storage.

In this Master's Thesis, we initially study the problem of WWW object storage in a WWW proxy's disks and analyse algorithms proposed for its solution. Resolving at first the problems caused by the creation and deletion of many small files, and secondly by the non-sequential disk accesses, these algorithms manage to enhance the throughput of storing WWW objects on disks by several times.

The main focus of this Thesis is the implementation of these algorithms for WWW object disk storage on a real WWW proxy, in order to evaluate them in a real-world situation. Accordingly, a WWW proxy system was developed from scratch, implemented with respect to the highest possible performance on a traditional Unix Operating System and the general-purpose Filesystem that accompanies it. The architecture of our prototype uses the asynchronous operations of TCP/IP sockets and a finite state machine (FSM), in order to process several WWW client requests in parallel.

The evaluation of our WWW Proxy prototype was conducted in a real-world setting using sophisticated WWW proxy benchmarks, like the Web Polygraph which is considered a de-facto standard for measuring the performance of commercial WWW proxies. Our experiments prove that our prototype has an average request service rate (throughput) more than three times higher than Squid, which is considered the most popular and advanced open source WWW proxy today.

Supervisor:

Evangelos P. Markatos
Associate Professor
Computer Science Department
University of Crete