Stepwell: Accurate Rate Limiting with Token Buckets for High-Speed Multi-Core Systems

Master's Thesis Project
ETH Zürich

Supervisors: Marc Wyss, Jelte van Bommel, Prof. Adrian Perrig
Project Duration: 6 months

Project Description

In high-performance networking and distributed systems, rate limiting is a critical mechanism to control traffic flow, prevent resource exhaustion, and ensure fairness among multiple consumers. The traditional token bucket (TB) algorithm is a widely used rate-limiting technique due to its simplicity and effectiveness. However, with more and more applications requiring rate limiting capabilities at very high speeds, the most simple solution might not be the most suitable one. Indeed, this is where traditional rate-limiting techniques become somewhat limited in performance, requiring expensive synchronization mechanisms.

This thesis will build upon a basic implementation of Stepwell, which introduced a hierarchical structure of TBs for rate limiting. By organizing rate limiters in a multi-level hierarchy, each core has access to its own dedicated TB, limiting the synchronization needs. The overall rate is still accurately enforced through the hierarchy, regardless of how the traffic is distributed across different processing cores. Stepwell is designed to work with non-locking TBs, addressing performance bottlenecks associated with locking and synchronization in traditional systems and making Stepwell particularly well-suited for multi-core environments.

In this thesis you will explore and analyze the space of existing rate-limiting mechanisms, implement them, and use these implementations to compare with Stepwell, but also improve and optimize Stepwell. The improved Stepwell makes use of a dedicated management thread to manage the TBs, making sure that each core does not require any atomic operations, and preventing frequent cache misses.

Tasks

Below are the tasks that the student is required to accomplish in the scope of this project. Based on the findings that the student makes and the issues he encounters during his work on the project, the goals of the project can later be changed and the tasks can be revised.

- Conduct literature review on high-speed rate limiting applications and usecases. Specifically look at:

1) other related work and their limitations, then reimplementing them in a common benchmark framework, and
2) the performance of TBs in standard libraries for common programming languages.

- Reimplement the Stepwell implementation in C/C++ and optimize on a low level to prevent pointer chasing, add prefetching, optimize commonly used paths and investigate the potential for SIMD.
- Find and implement a TB that is optimal in terms of memory storage.
- Try to avoid the compare-and-swap (CAS) spin lock in the atomic TB, e.g. by creating a mechanism where the TB provides more information with a request, such as the time to retry again.
- Create a synergistic approach that combines Stepwell and the atomic TB, switching between the two at different rates.
- Evaluate the different rate limiting approaches at different request rates and core counts, creating graphs that showcase when to switch to a different approach.

Optional Tasks

- Prove the correctness of the atomic TB, create an execution where a particular core is starved.
- Research the possibility for offloading rate limiting to the end host, using for instance Verifiable Delay Functions (VDFs).
- Research different fairness notions for TBs and how you can enforce these different notions.

Organization

The student will hold weekly meetings with the thesis advisor(s). During each weekly meeting, the student will be expected to briefly describe the work completed during the week. The student should promptly discuss any complications that arise (e.g. difficulty in understanding concepts or in creating tools) such that the advisor can assist the student in identifying alternative project directions. The advisor will assist the student toward completing any agreed upon milestones, as well as laying out the following week's goals.

Grading Scheme

Grade	**Description**
6.00	Design and implementation, as well as thesis are candidates for submission to an academic conference or workshop.
5.50	Thesis quality significantly exceeds expectations.
5.00	Thesis meets expectations.

4.50	Thesis partially meets expectations and has minor deficits.
4.00	Thesis meets minimum quality requirements; but has major deficits and is clearly below expectations