Efficient Synchronization of Linux Memory Regions over a Network: A Comparative Study and Implementation (Notes)

A user-friendly approach to application-agnostic state synchronization

Felicitas Pojtinger (Stuttgart Media University)

2023-08-04

Unsorted Research Questions

Structure

Structure

- Introduction
  - How does memory in Linux work? Paging, swap etc.
  - Examining Linux's memory management and relevant APIs
  - An introduction to mmap and how we can use it to map a file into memory/a byte slice, the role of msync
  - Concept: mmap a memory region with MMAP_SHARED to track changes in a file
  - Use cases for memory region synchronization (rough overview, esp. what is currently popular in the industry)
- Synchronization Strategies
  - **Implementing push-based memory sync by tracking changes to a mmaped slice with polled hashing of individual chunks**, why we can't use `inotify`, and the CPU-bound limitations of this approach
    - Detecting file changes: inotify; but mmap does not generate WRITE events
    - Comparing hashes of local and remote mmaped regions
    - Evaluation of hashing algorithms
    - Introduction to delta synchronization (e.g. rsync)