

Optimize Page Placement in Tiered Memory System

Thursday, September 23, 2021 7:00 AM (45 minutes)

Traditionally, all RAM is DRAM. Some DRAM might be closer/faster than others, but a byte of media has about the same cost whether it is close or far. But, with new memory tiers such as High-Bandwidth Memory or Persistent Memory, there is a choice between fast/expensive and slow/cheap.

We use the existing reclaim mechanisms for moving cold data out of fast/expensive tiers. It works well for that. However, reclaim does not work well for moving hot data which might be stuck in a slow tier since the pages near the top of the LRU are the most recently accessed only if there's regular memory pressure on the slow/cheap tiers.

Fortunately, NUMA Balancing *can* find recently-accessed pages regardless of memory pressure. We have repurposed it from being used for location-based optimization to being used for tier-based optimization. We have also optimized it for better hot data identification, such as to find frequently-accessed pages instead of recently-accessed pages, etc.

We will show our findings so far, and discuss the remaining problems, potential solutions, and alternatives.

The patchset email threads are as follows,

<https://lore.kernel.org/linux-mm/20210625073204.1005986-1-ying.huang@intel.com/>

<https://lore.kernel.org/linux-mm/20210311081821.138467-1-ying.huang@intel.com/>

I agree to abide by the anti-harassment policy

I agree

Primary author: HUANG, Ying

Presenter: HUANG, Ying

Session Classification: Performance and Scalability MC

Track Classification: Performance and Scalability MC