

Linux Plumbers Conference 2024



Contribution ID: 221

Type: **not specified**

Madvised lazy free

Friday, 20 September 2024 16:00 (15 minutes)

Conventional wisdom has held that madvised overhead has been mostly the syscall overhead. However, profiling shows this not to be the case.

Even on a medium sized 1 socket system, about half the CPU time spent in MADV_DONTNEED is spent flushing the TLB, and that is just in the calling CPU. Add in handling of the TLB flush IPIs on the other CPUs, and 90-95% of the MADV_DONTNEED overhead is TLB flushes.

I would like to propose MADV_LAZY_FREE, which can avoid most of the TLB flush overhead through collaborative behavior with the malloc library, and some inspiration from RCU.

Primary author: VAN RIEL, Rik (Facebook)

Presenter: VAN RIEL, Rik (Facebook)

Session Classification: Kernel Memory Management MC

Track Classification: Kernel Memory Management MC