



Contribution ID: 211

Type: **not specified**

## Reducing the number of users of mmap\_sem

*Wednesday, November 14, 2018 12:00 PM (15 minutes)*

The mmap\_sem has long been a contention point in the memory management subsystem. In this session some mmap\_sem related topics will be discussed. Some optimization has been merged by the upstream kernel to solve holding mmap\_sem for write for excessive period of time in munmap path by downgrading write mmap\_sem to read. And, some optimization are under discussion on the mailing list, i.e. release mmap\_sem earlier for page cache readahead, speculative page fault. There is still optimization room by figuring out just what mmap\_sem protects. It covers access to many fields in the mm\_struct structure. It is also used for the virtual memory area (VMA) red-black tree, the process VMA list, and various fields within the VMA structure itself. Finer grain locks might be better to replace mmap\_sem to reduce contention, i.e. range lock or per vma lock.

**Presenter:** SHI, Yang (Alibaba Group)

**Session Classification:** Performance and Scalability MC