IO 调度优化与 Debug

Cauchy Pu

August 25, 2020

Abstract

本文介绍内核的 IO 调度，利用 FIO 工具来测试 IO 性能，并以此工具来复现 IO 相关的 Bug 并最终基于对内核 IO 调度的分析来解决这个 Bug，同时尝试优化申威内核的 IO 调度。

Contents

Obviously the amount of time depends on where the head was previously located and how fortunate you are with the location of the sector on the platter: if it' s directly under the head you do not need to wait, but if it just passed the head you have to wait for a complete revolution. Even on the fastest 15k RPM disk that takes 4 milliseconds (15,000 rotations per minute = 250 rotations per second, which means one rotation is 1/250th of a second or 4ms). Admittedly that' s faster than the sushi in my earlier analogy, but the chances are you will need to read or write a far larger number of blocks than I can eat sushi dishes (and trust me, on a good day I can pack a fair few away).

What about the next block? Well, if that next block is somewhere else on the disk, you will need to incur the same penalties of seek time and rotational latency. We call this type of operation a random I/O. But if the next block happened to be located directly after the previous one on the same track, the disk head would encounter it immediately afterwards, incurring no wait time (i.e. no latency). This, of course, is a sequential I/O.