

Pmfs中目录项索引的实现

杨顺 国防科技大学计算机学院 (792454122@qq.com)

摘要

可字节寻址的非易失存储介质，如相变存储器，3D XPoint等，使数据可以在内存级别持久化。由于NVM本身的读写延时非常低，系统软件开销成为了决定整个持久化内存系统性能的主要因素。Pmfs是一个专门为持久化内存所设计的文件系统，它通过绕开页高速缓存，通用块层和块设备驱动来减少系统软件开销。然而，在Pmfs中每个目录操作(打开，创建或删除)都会遍历目录下的所有目录项，导致了随文件数增长而线性增长的目录项查找开销。通过测试发现，在特定类型负载下这种开销成为了整个文件系统的瓶颈。为解决该问题，本文在Pmfs中增加了持久化的目录项索引来加速目录操作。虽然这种方法带来了额外的一致性开销，但通过测试发现，基于单目录下100K文件的负载，该优化使得文件创建速度提高了12x，带宽增加了27.3%。

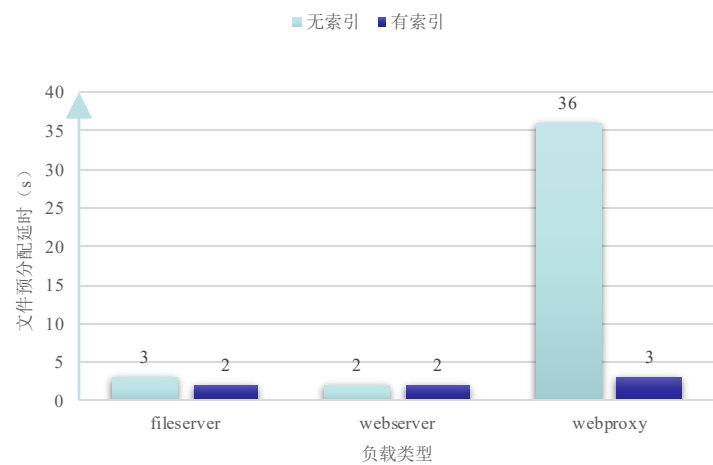
实验结果

负载名称	文件平均长度	目录宽度	文件数	进程数	读写比例	读写粒度
fileserver	128K	20	100K	8	1:2	1M/16K
webserver	16K	200	100K	8	10:1	1M/16K
Webproxy	16K	200K	100K	8	5:1	1M/16K

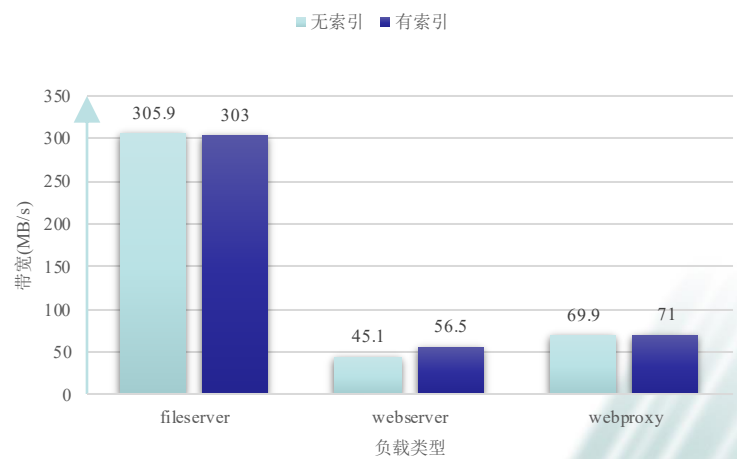
表1 filebench中三种不同类型负载的特点描述

负载编号	文件平均长度	目录宽度	文件数	进程数	读写比例	读写粒度
自定义一	32K	20	100K	8	5:1	1M/16K
自定义二	32K	200	100K	8	5:1	1M/16K
自定义三	32K	200K	100K	8	5:1	1M/16K
自定义四	128K	20	100K	8	5:1	1M/16K
自定义五	128K	200	100K	8	5:1	1M/16K
自定义六	128K	200K	100K	8	5:1	1M/16K

表2 自定义不同类型负载的特点描述

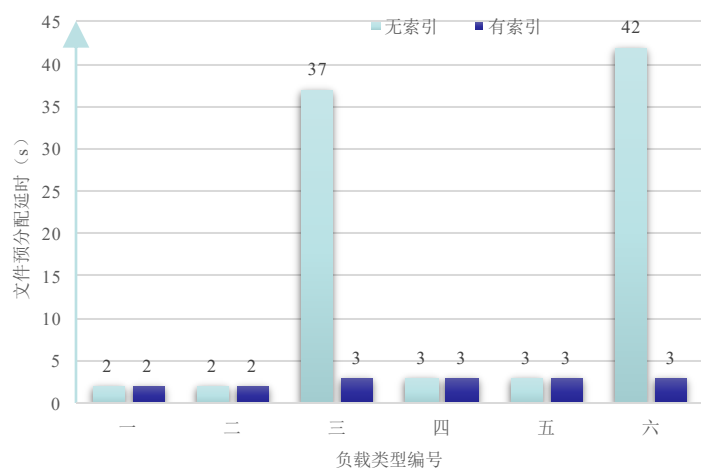


(a)文件预分配耗时统计图

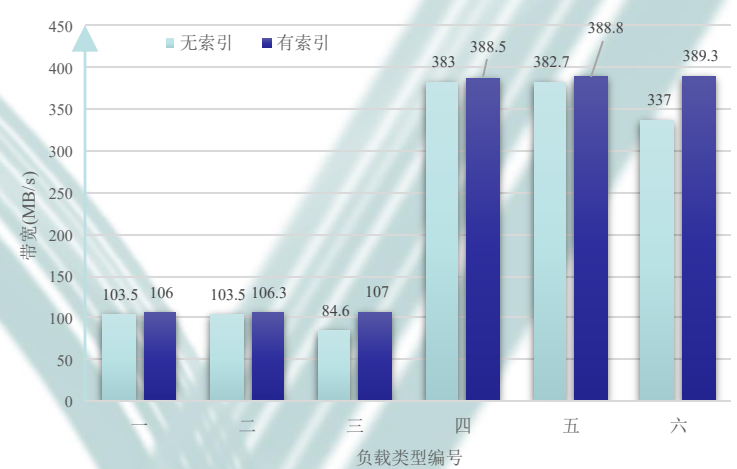


(b)带宽统计图

图1. 表一负载下pmfs增加目录项索引前后性能对比图



(a)文件预分配耗时统计图



(b)带宽统计图

图2.自定义负载下pmfs增加目录项索引前后性能对比图

总结

本文通过测试发现在单目录下文件量很大时，Pmfs的性能急剧下降，通过添加持久化的目录项索引，在带来了一定一致性开销的情况下，消除了目录项查找这一性能瓶颈。该优化适用于所有在Pmfs上进行二次开发的系统。

然而，在测试过程中发现，Pmfs还有很多地方存在性能瓶颈，如块分配器锁粒度太大导致并发写时的性能瓶颈。接下来，将不断地对Pmfs进行优化。只有这样经过多方优化后的Pmfs，在与传统文件系统对比时，才能体现出非易失内这种架构上的变化所带来的性能影响。