



Exploiting the performance gains of modern disk drives by enhancing data locality

Yuhui Deng

Cambridge-Cranfield High Performance Computing Facilities, Cranfield University Campus, Bedfordshire MK430AL, United Kingdom

ARTICLE INFO

Article history:

Received 19 September 2007

Received in revised form 22 January 2009

Accepted 1 February 2009

Keywords:

Disk drive
Data locality
Data access pattern
Block correlation
Data migration
Performance

ABSTRACT

Due to the widening performance gap between RAM and disk drives, a large number of I/O optimization methods have been proposed and designed to alleviate the impact of this gap. One of the most effective approaches of improving disk access performance is enhancing data locality. This is because the method could increase the hit ratio of disk cache and reduce the seek time and rotational latency. Disk drives have experienced dramatic development since the first disk drive was announced in 1956. This paper investigates some important characteristics of modern disk drives. Based on the characteristics and the observation that data access on disk drives is highly skewed, the frequently accessed data blocks and the correlated data blocks are clustered into objects and moved to the outer zones of a modern disk drive. The idea attempts to enhance spatial locality, improve the efficiency of aggressive sequential prefetch, and take advantage of Zoned Bit Recording (ZBR). An experimental simulation is employed to investigate the performance gains generated by the enhanced data locality. The performance gains are analyzed by breaking down the disk access time into seek time, rotational latency, data transfer time, and hit ratio of the disk cache. Experimental results provide useful insights into the performance behaviours of a modern disk drive with enhanced data locality.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

The storage hierarchy in current computer architectures is designed to take advantage of data access locality to improve overall performance. Each level of the hierarchy has higher speed, lower latency, and smaller size than lower levels. For decades, the hierarchical arrangement has suffered from significant bandwidth and latency gaps among processor, RAM, and disk drive [27,32,38]. The performance gap between processor and RAM has been alleviated by fast cache memories. However, the performance gap of RAM to disk drive has been widened to six orders of magnitude in 2000 and will continue to widen by about 50% per year [38].

Since the first disk drive was announced in 1956, disk drives have grown by over six orders of magnitude in density and over four orders in performance [29]. Over the last decade, areal density, track density and linear density have achieved 100%, 50%, and 30% growth respectively. Revolutions Per Minute (RPM) has been increased from 3600 in 1981 to 15,000 in 2000. Due to the significant growth in both the linear density and RPM, Internal Data Rate (IDR) has been growing at an exponential rate of 40% each year over the past 15 years [11]. Unfortunately, the basic mechanical architecture in disk drive has not changed too much. Slowed by the mechanical delays, disk access time was improved only about 8% per year [12]. Therefore, the disk I/O subsystem is repeatedly identified as a major bottleneck to system performance in many computing systems. The widening gap will be more serious when disk drives reach its physical limits due to the super paramagnetic effect.

E-mail addresses: deng_derek@emc.com, yuhuid@hotmail.com

The increasing performance gap between RAM and disk drive has long been a primary obstacle to improve overall system performance. To alleviate the impact of this widening gap, a lot of research efforts have been invested to improve disk access time. We only mention some of them which are related to our work in this paper. Ruemmler and Wilkes [36] employed disk shuffling to move frequently accessed data into the centre of a disk drive and organize the data into an organ pipe to reduce mean seek distances. They constructed a repeatable simulation environment across a range of workloads and disk drive types for comparing different shuffling algorithms. Their research indicated that the benefits are small to moderate, but are likely to be much larger with file systems that do not do a good initial data placement. Akyürek and Salem [1] presented an adaptive technique to copy a small number of frequently referenced disk blocks from their original locations to a reserved space near the middle of the disk. Their experiments showed that seek times are reduced 30–85% (depending on workloads) by adaptively rearranging about 3% of the data on the disk drive. FS2 [16] dynamically places multiple copies of data in file system's free blocks according to the disk access patterns observed at runtime to reduce the head positioning latencies. Because one or more replicas can be accessed in addition to their original data block, choosing the nearest replica that provides the fastest access can significantly improve disk I/O performance.

The above methods can substantially reduce seek time or rotational latency. However, those approaches may not be effective on modern disk drives due to evolutionary disk drive technologies. First of all, disk access time mainly consists of seek time, rotational latency and data transfer time. Due to the advance of Voice Coil Motors (VCM) electric drive and the increasing track density, long distance seeks of modern disk drives may not involve enormously more overhead than short ones. This results in a significant proportion decrease of seek time in disk access time. Secondly, due to geometric features, outer tracks on disk platters are much larger than the inner tracks. Modern disk drives employ a technique called Zoned Bit Recording (ZBR) to take advantage of its geometric features to increase disk capacity by varying the number of sectors per track with the distance from the spindle [30]. This characteristic results in much higher data transfer rate of outer zones than that of inner zones. Finally, the organ pipe is formed by placing the most frequently accessed cylinder in the middle of the disk drive, the next most frequently accessed cylinders on either side of the middle cylinder, and so on. This arrangement is provably optimal for independent disk accesses [36]. However, block correlations are common semantic patterns in storage systems, so the organ pipe arrangement could destroy the original block correlations [21]. The combination of these three reasons significantly counteracts the performance improvement of data reorganization.

This paper explores the performance gains of data reorganization based on a modern disk drive. Some new characteristics of modern disk drives, which are related to data reorganization, are reviewed. Based on the characteristics, blocks, which are correlated to the frequently accessed block, are clustered into objects, and the objects are moved to the outer zones of the disk drive. The aggressive sequential prefetch of the modern disk drive is enhanced and the block correlations remain due to the frequency based objects. Disk access time is broken into four basic components: seek time, rotational latency, data transfer time, and hit ratio of disk cache, each one is analyzed separately to determine its actual performance gains. Experimental results provide useful insights into the performance behaviour of block reorganization of a modern disk drive.

The remainder of this paper is organized as follows. An overview of modern disk drives is introduced in Section 2. Section 3 describes some important features of workload. The motivations of this paper are depicted in Section 4. Section 5 illustrates how to implement the block reorganization. The simulation environment and experimental validation are depicted in Section 6. Section 7 concludes the paper with remarks on main contributions and indications.

2. Modern disk drive overview

Disk access time T_{access} is mainly composed of seek time T_{seek} , rotational latency T_{rotate} and data transfer time $T_{transfer}$. The seek time measures the time for the disk head to move to a specified track. When the disk head arrives at the required track, the time spent on rotating the required sector to appear underneath the disk head is called rotational latency. The data transfer time is the amount of data divided by data transfer rate. T_{access} is expressed as follows:

$$T_{access} = T_{seek} + T_{rotate} + T_{transfer} \quad (1)$$

2.1. Seek time

Disk head is driven by a VCM to move over the recording surface to seek a target sector. In order to reduce the heat dissipation of the VCM, the temperature of the coil in the VCM is controlled by selecting a fixed maximum current for the seek distances which exceed a threshold. This threshold is typically 35% of a full stroke [40]. For a long seek distance which exceeds the threshold, the current in the coil reaches the maximum value when the disk head reaches a nominal maximum velocity. At the end of an acceleration period, the current is removed from the coil, which incurs a coast period that maintains a nominal maximum velocity. Then, the fixed maximum current is applied to the coil in an opposite direction to decelerate the disk head. When the disk head reaches the target track, a procedure is triggered to verify the current position. The time cost for the disk head to settle at the end of a seek is called settling time. Therefore, a seek time is composed of an acceleration time t_{acc} , a coast period t_{coast} , a deceleration time t_{dec} , and a head settling time t_{settle} . The acceleration time and deceleration time are proportional to the square root of the seek distance. The coast period is linear in the seek distance. For a short seek (e.g. single cylinder seek), the disk arm accelerates and decelerates without reaching the nominal maximum velocity.

Download English Version:

<https://daneshyari.com/en/article/395663>

Download Persian Version:

<https://daneshyari.com/article/395663>

[Daneshyari.com](https://daneshyari.com)