

各种提高文件系统性能的方法

# 文件系统的性能3

下面呢我们继续介绍提供文件系统性能的各种方法。



# 信息的优化分布

记录在磁道上的排列方式也会影响输入输出操作的时间

例子：

处理程序要求顺序处理8个记录；磁盘旋转一周为20毫秒/周；花5毫秒对记录进行处理



而前面那个呢 花的时间要多得多。





# 记录的成组与分解

## ➤ 记录的成组

把若干个逻辑记录合成一组存放一块的工作

## ➤ 进行成组操作时必须使用内存缓冲区，缓冲区的长度等于逻辑记录长度乘以成组的块因子

## ➤ 成组目的：提高了存储空间的利用率；减少了启动外设的次数，提高系统的工作效率

## ➤ 记录的分解

从一组逻辑记录中把一个逻辑记录分离出来

典型例子—  
目录文件





# RAID技术

美国加州伯克利分校  
D.A.Patterson教授  
1988年提出

**RAID** (独立磁盘冗余阵列)

(Redundant Arrays of Independent Disks)

多块磁盘按照一定要求构成一个独立的存储设备

目标: 提高可靠性和性能

考虑: 磁盘存储系统的速度、容量、容错、数据

灾难发生后的数据恢复

还包括了容错, 以及当出现了数据灾难之后 进行恢复的问题。





# RAID技术的思想

## 基本思路

数据是如何组织的？

- 通过把多个磁盘组织在一起，作为一个逻辑卷提供磁盘跨越功能
- 通过把数据分成多个数据块，并行写入/读出多个磁盘，以提高数据传输率（数据分条stripe）
- 通过镜像或校验操作，提供容错能力（冗余）

最简单的RAID组织方式：镜像

最复杂的RAID组织方式：块交错校验





# RAID 0 - 条带化

无冗余(即  
无差错控制)  
性能最佳

- 数据分布在阵列的所有磁盘上
- 有数据请求时，同时多个磁盘并行操作
- 充分利用总线带宽，数据吞吐率提高，驱动器负载均衡



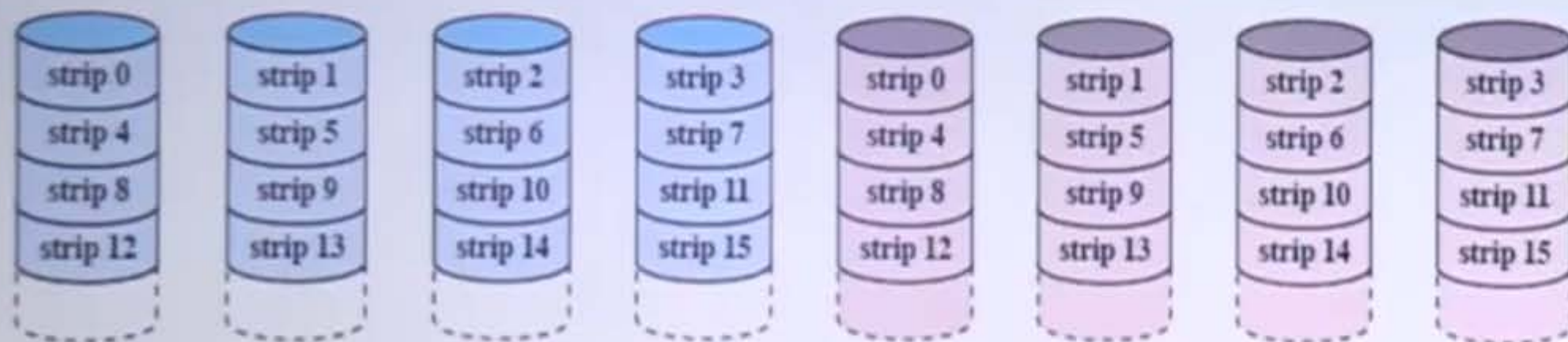
差错控制，但是性能是最快的 那么 RAID 1



# RAID 1 - 镜像

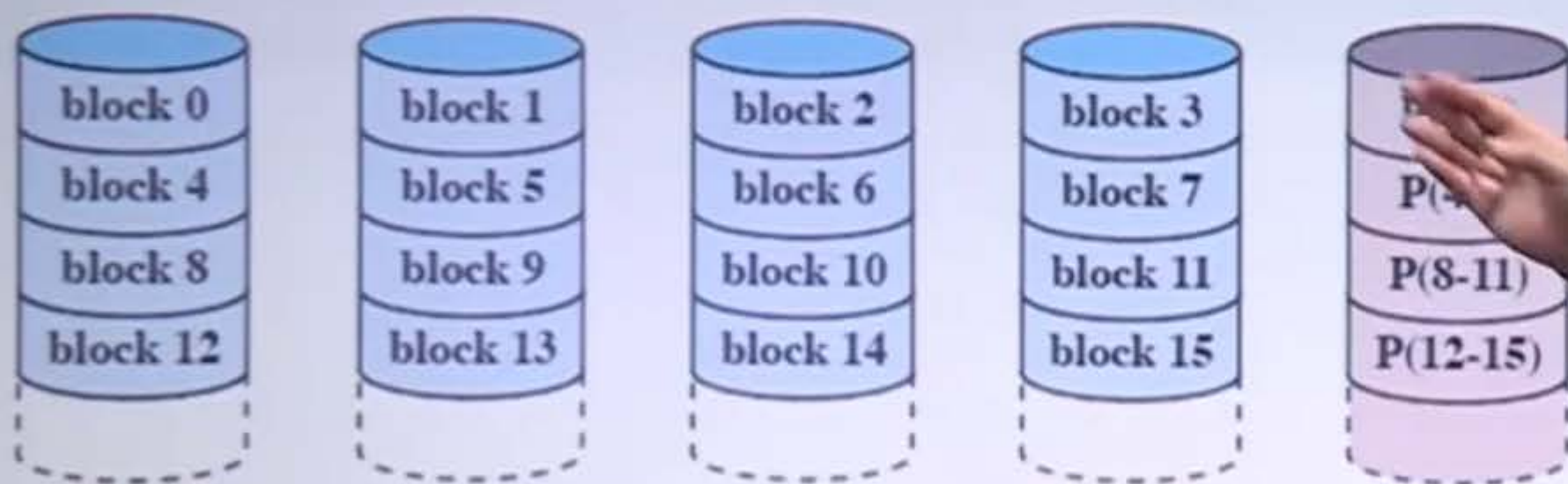
数据安全性  
最好

- 最大限度保证数据安全及可恢复性
- 所有数据同时存在于两块磁盘的相同位置
- 磁盘利用率**50%**



# RAID 4 交错块奇偶校验

- 带奇偶校验
- 以数据块为单位





下面呢我们继续介绍提供文件系统性能的各种方法。下面我们要介绍的呢是信息的优化分布 它指的是说如果 我们的信息是按记录划分, 记录在 磁道上的排列方式它会去影响输入输出操作的时间 比如说我们给一个例子, 这个例子呢 我们有一个程序, 这个程序呢要顺序处理 8 个记录 那么磁盘旋转一圈呢, 假设是 20 毫秒, 需要花 20 毫秒旋转一圈 那也就是说这 8 个记录呢, 每个记录呢 2.5 毫秒就要把这个记录读完了 读完一个记录之后呢, 还要花 5 毫秒对这个记录进行处理, 然后再读第二个记录 假设我们编程序是先读一个记录, 处理, 再读一个记录, 再处理 假设我们的这个记录正好是按 这样一个方式分布的, 那么当我读第一个记录的时候呢 需要花 2.5 毫秒, 然后去花 5 毫秒去处理 等处理完了之后, 这时候, 因为有旋转 延迟, 所以呢已经是 4 号扇区, 4 号记录 经过我的磁头了。那么我要 读第二个记录, 我就得需要等到 4,5,6,7,8,1 都经过之后才能够读 2 号记录 那么我们可以看一下它这个顺序, 按照这样一个顺序的话, 我们读第 2 个记录的话, 需要跳过 4,5,6,7,8,1, 那么就需要跳过 2.5 毫秒乘以 6, 我们得需要跳过这么长时间才能去读 2 号记录, 然后花 2.5 毫秒去读 然后再花 5 毫秒去处理, 所以可以算出来对 2 号记录的这个处理需要花  $2.5 * 6$ , 再加上这个 2.5, 再加上 5, 这么多个时间 那么如果我们对信息进行一个优化分布呢 比如说我们按这样一个方式来 安排记录的位置, 比如说 1 号记录 之后呢我是放 4 号记录, 4 号记录放 7 号记录, 然后再放 2 号记录 因此我们来看, 如果我读完 1 号记录, 又花了 5 毫秒去处理完了之后 正好 2 号记录经过磁头, 可以继续去读, 所以呢总的时间呢 就会降低, 因为每个记录处理的时间呢是 2.5 毫秒加上 5 毫秒, 是 7.5 毫秒 一共要处理 8 个记录呢, 乘以 8 就可以了。而前面那个呢 花的时间要多得多。还有一种 提高文件系统性能的方法呢我们称之为记录的成组与分解 所谓记录的成组是说把若干个记录合并成一组存放在一个数据块里头 因为如果我们的文件是按记录这样一个格式 那么是一个记录就占一块呢, 还是把若干记录 组织在一起, 当然了, 是整数倍, 那么就放在同一个块里头 那我们这里就讨论的是记录的成组, 用这种方式可以提高文件系统性能 在进行这个成组操作的时候呢, 必须有一个内部缓冲区 缓冲区的长度呢实际上是 记录的长度乘以成组的块因子, 也就是说 一块里头可以放多少个记录, 这称之为块因子, 所以我们可以定义一个 缓冲区的长度, 是记录的长度乘以块因子, 得到的这样一个缓冲区的大小 我们通过这个缓冲区呢, 主要的目的 成组的主要的目的是为了提高 存储空间的利用率, 不是一个记录占一块, 而是若干记录占一块 减少了启动外部设备的次数, 因为我们如果一个记录启动一次 和



我把若干个记录放到一块，启动一次，那显然减少了这个次数 当然了，目的是为了整个文件系统的性能 有了成组就有另外一个，叫分解 因为我读的时候也是按块读，所以把一组记录读出来，我要从中间找到我所需要的那一个记录，那么这就就是说，把那个记录把它分离出来，这叫记录的分解 当然这种方式呢比较容易想到，因为我们 典型的一个例子就是目录文件，其实目录文件就是按这样一个记录的成组方式 来完成它的这个实现的，分配和这个使用的过程的 下面我们介绍一下 RAID 技术 RAID 呢是指独立的磁盘冗余阵列 它是 1988 年 加州伯克利分校 Patterson 教授提出来的一种技术 它的设计思想非常简单，就是 将多块磁盘按照一定的要求 构造成一个独立的存储设备 目标呢就是为了提高 整个存储设备的可靠性和性能 因此，RAID 技术在实现的时候 考虑到了磁盘存储系统的方方面面 包括了速度、性能方面，包括容量方面 还包括了容错，以及当出现了数据灾难之后 进行恢复的问题。当把多块盘 按照一定的要求给它构造成一个独立的存储设备时 那么对于文件系统而言，数据是怎么组织的呢？我们来看一下 RAID 技术的基本思想 首先呢，是通过把多个磁盘组织在一起 作为一个逻辑卷 提供给文件系统，让这个文件系统实际上是有一个跨盘的功能 第二个呢是保存文件的时候 把文件的数据划分成多个数据块 然后并行写入或者是读出多个磁盘 因为我有多块盘，我可以把数据块 分布在多个盘上，所以在读写盘的时候是一个并行的动作 因此可以提高数据传输的速率 那么这种工作呢我们就称之为数据分条 因为要跨盘存放，把数据分条之后，每一条在一块盘上 第三种就是要通过镜像 或者是校验，提供这些手段 来提高这样一个存储设备的这个容错能力 当然这就是说通过了一些冗余的信息的保存 能够提高容错能力。RAID 的组织方式呢有多种 最简单的 RAID 组织方式呢称之为镜像 而最复杂的 RAID 组织方式呢称为块交错校验 我们这里举几个例子。第一种呢叫 RAID 0 RAID 0 呢实际上就是一个条带化 它把这个数据，一个文件的数据是分布在这个 磁盘阵列当中的所有盘上 所以当有数据请求来的时候，可以多个磁盘并行操作 这种方式呢我们来看一个例子 我们把一个文件的数据给它 分条，划分成若干个分条，条带化 然后我们假设有 4 块盘，我们就可以 0,1,2,3 同时写入盘上 然后再 4,5,6,7 再同时写入，读的时候也可以这样去做 所以一次就可以保存更多的信息在磁盘上 所以这种方法呢是充分利用了总线的带宽，数据的吞吐率非常高 当然了，驱动器的负载也是均衡的。那么在 RAID 0 呢它的主体思想呢是条带化，所以它没有冗余的信息存放，



也就是说它没有 差错控制，但是性能是最快的 那么 RAID 1 呢就是简单的一种镜像 它的设计思想是最大限度地保证数据安全及可恢复性 因此呢，我不管有多少块盘，那么我的所有的 数据都同时保存在两块盘的相同位置 也就是使得磁盘的利用率达到的是 50% 我们来看一下，同样，刚才的例子，那我们就要 8 块盘 8 块盘呢，那么前面 4 个条带呢放在前面 4 块盘的这个位置，那么同样，镜像，镜像 盘呢也是把这 4 个条带呢放在相同的这个位置 所以它的数据安全性是最好的，但是呢磁盘的利用率呢是不高的 那我们来看一个，另外一个 RAID 4，其实还有很多的 RAID，我们就举 RAID 4 做个例子 RAID 4 呢叫做交错块奇偶校验 那么首先它的校验，也就是说来提升可靠性的呢是通过一个校验 来完成，而这个校验呢是奇偶校验 所以它是奇偶校验的一个信息，要存放在磁盘上 而且呢，它的单位，划分单位就是数据块，数据块我们再来看这个例子。 如果一个文件我们就按块大小划分成若干块 那假设有 4 块盘，有 5 块盘，其中呢数据部分呢按块 分别保存在 4 块盘上，它的 奇偶校验的信息呢是放在了第 5 块盘上 因此如果读的时候发现了一些错误的话，我们可以 通过第 5 块盘上的校验的信息呢来进行相应的恢复 而把一个文件的数据 分条，按块来分条，分别放在不同的磁盘上呢，也 能够提高性能，所以它是一个比较综合的，既提高了性能 又能够保证一定的可靠性，它是这样一样