(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 103530069 A
(43)申请公布日 2014.01.22

(21)申请号 201310542533.8

(22)申请日 2013.11.05

(71)申请人 浪潮(北京)电子信息产业有限公司
地址 100085 北京市海淀区上地信息路2号2-1号C栋1层

(72)发明人 周耀辉 李伟国

(74)专利代理机构 北京安信方达知识产权代理有限公司 11262
代理人 王丹 栗若木

(51)Int.Cl.
G06F 3/06(2006.01)

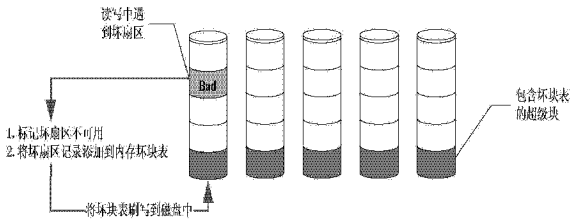
权利要求书1页 说明书2页 附图2页

(54)发明名称

一种RAID5磁盘阵列坏扇区处理方法

(57)摘要

提供一种RAID5磁盘阵列坏扇区处理方法，在创建RAID5磁盘阵列的时候，在每个RAID成员磁盘的超级块中增加4KB的坏块记录空间，用于记录每个成员磁盘中的坏扇区位置，当RAID5磁盘阵列启动的时候，将每个成员磁盘的坏块记录空间的数据映射到内存中，在内存中建立一个坏块表，在执行磁盘扇区写操作时以及在执行磁盘扇区读操作遇到错误时，根据所述内存坏块表中对该磁盘扇区的记录执行相应的操作。提出的所述方法在磁盘中出现坏扇区时不用剔除该磁盘，只是标记出坏扇区，对其不用读写，极大的提高RAID5磁盘阵列的稳定性，延长磁盘阵列使用寿命，提高存储资源利用率，保证数据可靠性。



CN 103530069 A

1. 一种 RAID5 磁盘阵列坏扇区处理方法，其特征在于包括：

S1：在创建 RAID5 磁盘阵列的时候，在每个 RAID 成员磁盘的超级块中增加 4KB 的坏块记录空间，用于记录每个成员磁盘中的坏扇区位置；

S2：当 RAID5 磁盘阵列启动的时候，将每个成员磁盘的坏块记录空间的数据映射到内存中，在内存中建立一个坏块表；

S3：在执行磁盘扇区写操作时，根据所述内存坏块表中对该磁盘扇区的记录执行相应的操作；

S4：在执行磁盘扇区读操作时，若遇到扇区读错误，则根据所述内存坏块表中对该磁盘扇区的记录执行相应的操作。

2. 如权利要求 1 所述的方法，其特征在于：

所述步骤 S3 中所述相应的操作具体为先在内存坏块表中查看所述扇区是否已经留有记录，如果确认有所述扇区的记录，就跳过此扇区，将数据写入别的扇区；如果内存坏块表中没有目标扇区的记录，则将数据写入该扇区。

3. 如权利要求 1 或 2 所述的方法，其特征在于：

所述步骤 S3 还包括，如果写操作失败，则在内存坏块表中添加当前扇区记录。

4. 如权利要求 3 所述的方法，其特征在于：

采用二分查找插入法添加当前扇区记录，并按照扇区的大小在内存坏块表中排序扇区。

5. 如权利要求 1 所述的方法，其特征在于：

所述步骤 S4 中所述相应的操作具体为：先在内存坏块表中查看目标扇区是否已经留有记录，如果有记录，跳过读取别的扇区；如果没有记录，则通过条带中其余磁盘的数据异或校验获得这块坏扇区上的数据。

6. 如权利要求 5 所述的方法，其特征在于：

将获得的数据重新写在发生读操作错误的扇区上，如果写操作失败就将在内存坏块表中添加该扇区的记录。

一种 RAID5 磁盘阵列坏扇区处理方法

技术领域

[0001] 本发明涉及磁盘阵列技术领域，具体涉及一种 RAID5 磁盘阵列坏扇区处理方法。

背景技术

[0002] RAID 技术就是利用容量比较小的磁盘按一定的规则组成一个大容量的磁盘阵列向外提供存储资源，RAID(Redundant Array of Independent Disks) 就被称为独立冗余磁盘阵列。

[0003] 使用 RAID 的主要目的有：

[0004] (1) 对磁盘上的数据进行冗余存储，实现容错功能。

[0005] (2) 有 I/O 请求时，可以对磁盘进行并发访问，提高吞吐率。

[0006] 如附图 1 所示，RAID5 支持冗余，在阵列中的所有磁盘上都存储检验信息，其校验和技术是在磁盘条带中采用异或校验，被称作为“采用块交叉访问及校验信息均匀分布”模式。条带是 RAID 成员盘的存储空间被划分为大小相等的小空间。异或运算特征为 A ⊕ B=C==>B ⊕ C=A，A ⊕ C=B。因此，如果 A、B 为数据块，C 为校验块，当 B 所在的盘失效后，能够通过 A 和 C 异或校验计算出 B。这是 RAID5 容错的最基本原理，可以容忍一块磁盘故障。

[0007] RAID5 读写操作是以条带为基本单位，条带的宽度为 4KB，条带的长度为成员磁盘个数。对于条带中磁盘数据的写操作，都要更新校验盘数据，以确保条带数据的一致性。

[0008] 目前 RAID5 读写操作遇到坏扇区的时候，将有坏扇区的磁盘从 RAID5 磁盘阵列中剔除，RAID5 降级，降级后的 RAID5 失去了容错能力，但是还可以进行读写操作，其读写操作如附图 2 和 3 所示。如果降级后的 RAID5 其成员盘再遇到坏扇区，会将此坏扇区的磁盘从 RAID5 阵列中剔除，由于 RAID5 掉两块盘，整个阵列失效，所有的数据会丢失。

[0009] 目前的 RAID5 技术在磁盘阵列中的两块磁盘中出现坏扇区时，就会导致整个阵列不可用，数据丢失，极大的浪费了存储资源，也不利于数据安全。

发明内容

[0010] 为了解决上述技术问题，本发明提出了一种 RAID5 磁盘阵列坏扇区处理方法，可以提高 RAID5 磁盘阵列的可靠性、稳定性，延长阵列的寿命，提高存储资源利用率。

[0011] 一种 RAID5 磁盘阵列坏扇区处理方法，包括：

[0012] S1：在创建 RAID5 磁盘阵列的时候，在每个 RAID 成员磁盘的超级块中增加 4KB 的坏块记录空间，用于记录每个成员磁盘中的坏扇区位置；

[0013] S2：当 RAID5 磁盘阵列启动的时候，将每个成员磁盘的坏块记录空间的数据映射到内存中，在内存中建立一个坏块表；

[0014] S3：在执行磁盘扇区写操作时，根据所述内存坏块表中对该磁盘扇区的记录执行相应的操作；

[0015] S4：在执行磁盘扇区读操作时，若遇到扇区读错误，则根据所述内存坏块表中对该

磁盘扇区的记录执行相应的操作。

[0016] 特别地，所述步骤 S3 中所述相应的操作具体为先在内存坏块表中查看所述扇区是否已经留有记录，如果确认有所述扇区的记录，就跳过此扇区，将数据写入别的扇区；如果内存坏块表中没有目标扇区的记录，则将数据写入该扇区。

[0017] 特别地，所述步骤 S4 中所述相应的操作具体为：先在内存坏块表中查看目标扇区是否已经留有记录，如果有记录，跳过读取别的扇区；如果没有记录，则通过条带中其余磁盘的数据异或校验获得这块坏扇区上的数据。

[0018] 本发明的有益效果是：采用本发明提出的 RAID5 坏扇区处理方法，在磁盘中出现坏扇区时不用剔除该磁盘，只是标记出坏扇区，对其不用读写，极大的提高 RAID5 磁盘阵列的稳定性，延长磁盘阵列使用寿命，提高存储资源利用率，保证数据可靠性。

附图说明

[0019] 图 1 是现有技术中的 RAID5 结构图。

[0020] 图 2 是现有技术中在 RAID5 降级后，目标扇区在坏盘上的读操作。

[0021] 图 3 是现有技术中在 RAID5 降级后，目标扇区在坏盘上的写操作。

[0022] 图 4 是本发明提出的 RAID5 读写遇坏扇区处理方法示意图。

具体实施方式

[0023] 下面参照附图 4，对本发明的内容以一个具体实例来描述本发明提供的所述方法。

[0024] 在创建 RAID5 磁盘阵列的时候，在每个 RAID 成员盘的超级块中增加 4KB 的坏块记录空间，用于记录每个成员盘中的坏扇区位置。每条坏块记录用 64bit 来表示，最高位表示这条记录是否已经确认，如果最高位置位了，说明这条记录的扇区为坏扇区不可用，反之不是；低 9 位为此条记录表示的坏扇区的大小；中间 54 位为坏扇区的起始地址。

[0025] 当 RAID5 磁盘阵列启动的时候，将每个成员磁盘坏块记录空间的数据映射到内存中，在内存中建立一个坏块表。当 RAID5 写操作的时候，先在内存坏块表中查看目标扇区是否已经留有记录，如果确认有目标扇区的记录，就跳过此扇区，将数据写入别的扇区。如果坏块表中没有目标扇区的记录，就说明此扇区正常，直接刷写。如果写扇区失败，则将这个扇区记录添加到内存坏块表中，标明该扇区为坏扇区不可用。添加坏扇区记录采用二分查找插入法，坏块记录按照坏块的扇区大小排序，可以提高查找检索效率。当 RAID5 读操作的时候，遇到扇区读错误，先在内存坏块表中查看目标扇区是否已经留有记录，如果有记录，跳过读取别的扇区；如果没有记录，说明这块扇区上有写数据，可以通过条带中其余磁盘的数据异或校验获得这块坏扇区上的数据，并且将获得的数据重新写在发生读操作错误的扇区上，如果写操作也失败就将此坏扇区添加到坏块表中。

[0026] RAID5 每个成员盘的内存坏块表会在 RAID5 更新超级块的时候，同超级块一同刷写到成员磁盘的坏块记录空间中。

[0027] 当然，本发明还可有其他多种实施例，在不背离本发明精神及其实质的情况下，熟悉本领域的技术人员当可根据本发明作出各种相应的改变和变形，但这些相应的改变和变形都应属于本发明的权利要求的保护范围。

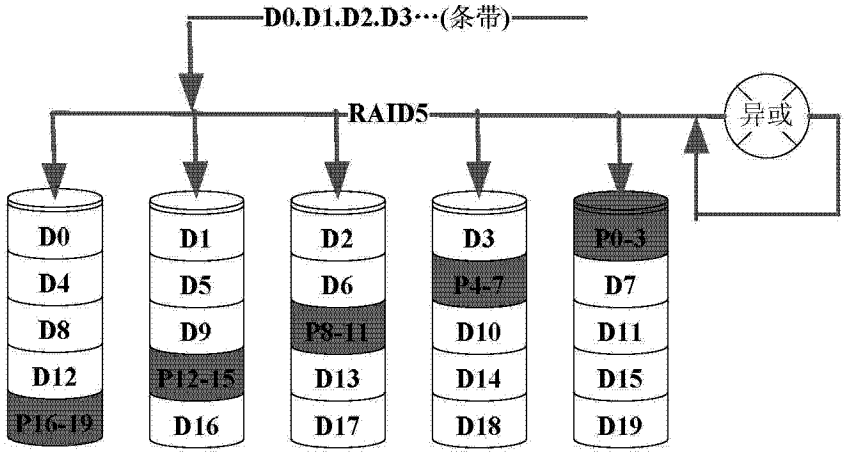

D0.D1.D2.D3…(条带)
RAID5
异或
D0
D4
D8
D12
P16-19
D1
D5
D9
P12-15
D16
D2
D6
P8-11
D13
D17
D3
P4-7
D10
D14
D18
P0-3
D7
D11
D15
D19

图 1

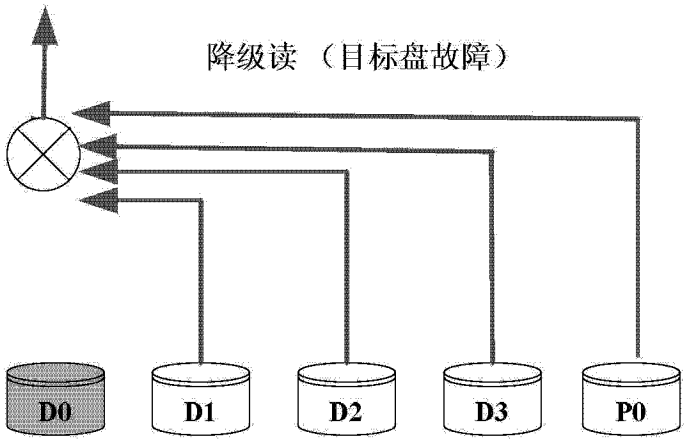

降级读 （目标盘故障）
D0
D1
D2
D3
P0

图 2

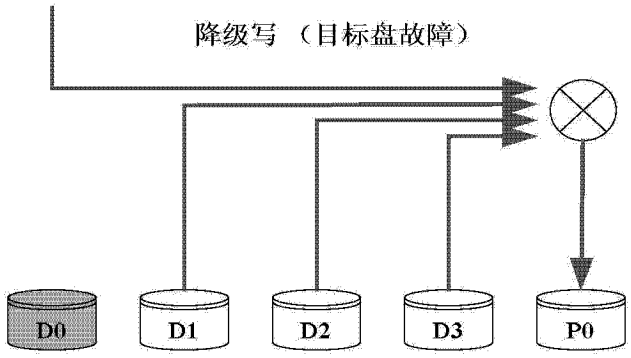

降级写 （目标盘故障）
D0
D1
D2
D3
P0

图 3

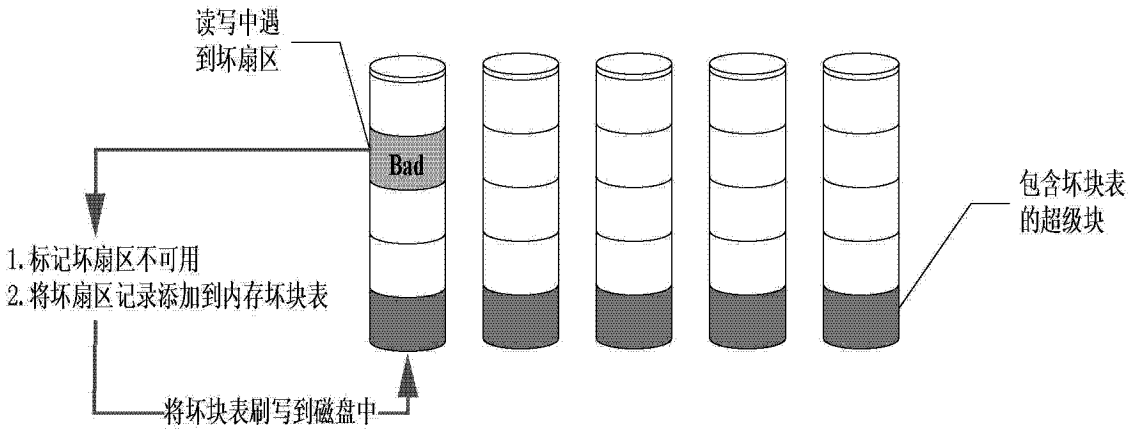

读写中遇
到坏扇区
Bad
包含坏块表
的超级块
1.标记坏扇区不可用
2.将坏扇区记录添加到内存坏块表
将坏块表刷写到磁盘中

图 4