



(12)发明专利

(10)授权公告号 CN 106502587 B

(45)授权公告日 2019.10.25

(21)申请号 201610912077.5

(22)申请日 2016.10.19

(65)同一申请的已公布的文献号
申请公布号 CN 106502587 A

(43)申请公布日 2017.03.15

(73)专利权人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 丁敬文

(74)专利代理机构 深圳市深佳知识产权代理事务所(普通合伙) 44285

代理人 王仲凯

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

(56)对比文件

CN 105956090 A,2016.09.21,
CN 103514260 A,2014.01.15,
CN 103544045 A,2014.01.29,
CN 105224237 A,2016.01.06,
US 2006174156 A1,2006.08.03,

审查员 张滔

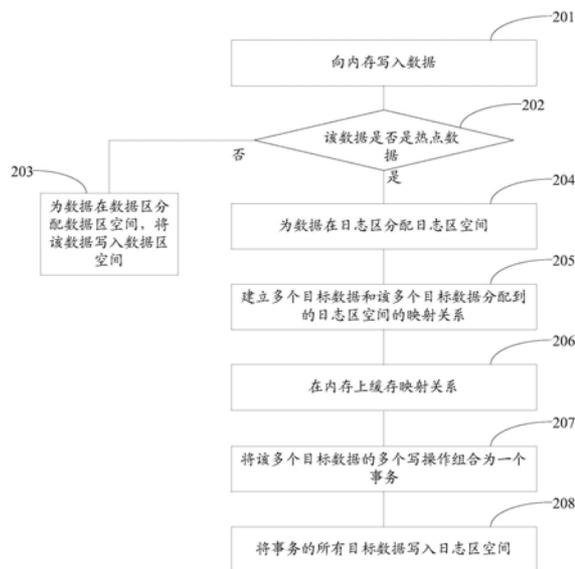
权利要求书5页 说明书25页 附图6页

(54)发明名称

硬盘数据管理方法和硬盘控制装置

(57)摘要

本发明实施例公开了种硬盘数据管理方法和硬盘控制装置,用于高效管理硬盘上的碎片。本发明实施例应用于包括硬盘的硬盘控制装置,该硬盘包括数据区和日志区,方法包括:向缓存器件写入数据;判断数据是否是热点数据,其中热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据;若数据不是热点数据,则为数据在数据区分配数据区空间,将数据写入数据区空间;若数据是热点数据,则为数据在日志区分配日志区空间,将数据写入日志区空间。通过在硬盘上将不同类型的数据保存在不同的区域以不同的方式进行管理,可提高硬盘上的碎片管理效率,日志区对硬盘碎片的高效管理,可减少硬盘碎片的产生。



1. 一种硬盘数据管理方法,其特征在于,所述方法应用于包括硬盘的硬盘控制装置,所述硬盘包括数据区和日志区,所述方法包括:

向缓存器件写入数据;

判断所述数据是否是热点数据,其中所述热点数据为存储在所述硬盘上后在预设次数的修改和释放后能使所述硬盘产生预设数量碎片的数据;

若所述数据不是热点数据,则为所述数据在所述数据区分配数据区空间,将所述数据写入所述数据区空间;

若所述数据是热点数据,则为所述数据在所述日志区分配日志区空间,将所述数据写入所述日志区空间;

所述方法还包括:

当所述数据区的空间利用率大于预设数据区利用阈值时,将当前空闲的日志区转化为数据区;

当所述日志区的空间利用率大于预设日志区利用阈值时,将由空闲的日志区转化成的数据区转化为日志区。

2. 根据权利要求1所述的方法,其特征在于,所述缓存器件为内存,

所述为所述数据在所述日志区分配日志区空间之后,所述方法还包括:

建立所述数据和所述日志区空间的映射关系。

3. 根据权利要求2所述的方法,其特征在于,

所述建立所述数据和所述日志区空间的映射关系,包括:

建立多个目标数据和所述多个目标数据分配到的日志区空间的映射关系,其中所述目标数据属于热点数据;

所述将所述数据写入所述日志区空间,包括:

将所述多个目标数据的多个写操作组合为一个事务;

将所述事务的所有目标数据写入日志区空间,当所述事务的其中一个目标数据的写操作执行失败时,所述事务的其他目标数据执行的写操作失败。

4. 根据权利要求3所述的方法,其特征在于,所述方法还包括:

在所述内存上缓存属于所述热点数据的数据。

5. 根据权利要求4所述的方法,其特征在于,

所述将所述事务的所有目标数据写入日志区空间之前,所述方法还包括:

根据所述多个目标数据建立数据链表,其中,所述数据链表用于管理所述目标数据,所述数据链表管理的目标数据与所述事务的目标数据相同;

根据所述数据链表对所述目标数据进行管理;

在预设释放条件下,根据所述数据链表从先到后的建立顺序,查找所述数据链表未解除管理的数据;

在所述内存上释放目标数据链表未解除管理的数据,且在所述内存上保留所述目标数据链表对应的目标映射关系;

其中,所述根据所述数据链表对所述目标数据进行管理,包括:

建立第二数据链表后,当所述第二数据链表管理的第二目标数据是由预先建立的第一数据链表管理的第一目标数据修改得到时,在所述第一数据链表上解除对所述第一目标数

据的管理；在与所述第一数据链表对应的第一映射关系上删除所述第一目标数据的信息。

6. 根据权利要求5所述的方法，其特征在于，

所述在预设释放条件下，根据所述数据链表从先到后的建立顺序，查找所述数据链表未解除管理的数据，包括：

当所述内存达到第一预设水位时，根据所述数据链表从先到后的建立顺序，查找所述数据链表未解除管理的数据；

所述在所述内存上释放目标数据链表未解除管理的数据之后，所述方法还包括：

在所述内存达到第二预设水位时，从所述日志区读取所述目标映射关系指向的数据；

将所述目标映射关系指向的数据写入所述数据区；

在所述内存上删除所述目标映射关系。

7. 根据权利要求5所述的方法，其特征在于，

所述方法还包括：

在预设回收条件下，执行日志区数据搬迁的步骤；

所述执行日志区数据搬迁的步骤，包括：

查找所述映射关系；

根据映射关系记录的信息判断与所述映射关系对应的第一日志区上的数据是否迁移完；

若所述第一日志区上的数据未迁移完，则根据所述映射关系记录的信息，确定第一日志区上的空间利用率；

当所述第一日志区的空间利用率小于预设利用率阈值时，将所述第一日志区的数据迁移至第二日志区，并更新与所述被搬迁的数据对应的映射关系，其中所述第二日志区为空闲的日志区或在回收日志区时使用过的日志区；

当当前日志区总的空间水位达到预设空间阈值时，则停止执行日志区数据搬迁的步骤，否则继续执行日志区数据搬迁的步骤。

8. 根据权利要求7所述的方法，其特征在于，

所述方法还包括：

根据所述事务的写入顺序为所述事务对应的数据链表按照递增规则分配事务号；

在所述执行日志区数据搬迁的步骤中，所述查找所述映射关系，包括：

从当前事务号最小的数据链表开始，根据所述事务号由小到大的顺序查找与所述数据链表对应的映射关系。

9. 根据权利要求4所述的方法，其特征在于，

所述在所述内存上缓存属于所述热点数据的数据，包括：

所述判断所述数据是否是热点数据之后，若所述数据是热点数据，则在所述内存上保留所述数据；或者，

所述向缓存器件写入数据之前，从所述日志区读取数据到所述缓存器件缓存。

10. 根据权利要求1至9任一项所述的方法，其特征在于，所述热点数据包括数据大小小于预设数据阈值的数据和/或所述热点数据包括元数据。

11. 根据权利要求1至9任一项所述的方法，其特征在于，

所述为所述数据在所述日志区分配日志区空间，将所述数据写入所述日志区空间，包

括：

为所述数据在所述日志区按照顺序分配日志区空间，将所述数据顺序追加写入所述日志区空间。

12. 根据权利要求2至9任一项所述的方法，其特征在于，

所述方法还包括：

为所述映射关系在日志区上分配日志区空间；

将所述映射关系写入所述映射关系分配到的日志区空间。

13. 一种硬盘控制装置，其特征在于，所述硬盘控制装置包括：

写入单元，用于向缓存器写入数据；

缓存管理器，用于判断所述数据是否是热点数据，其中所述热点数据为存储在所述硬盘上后，在预设次数的修改和释放后能使所述硬盘产生预设数量碎片的数据；

数据管理器，用于若所述数据不是热点数据，则为所述数据在硬盘的数据区分配数据区空间，将所述数据写入所述数据区空间；

日志管理器，用于若所述数据是热点数据，则为所述数据在所述硬盘的日志区分配日志区空间，将所述数据写入所述日志区空间；

所述硬盘控制装置还包括：

数据区转化单元，用于当所述数据区的空间利用率大于预设数据区利用阈值时，将当前空闲的日志区转化为数据区；

日志区转化单元，用于当所述日志区的空间利用率大于预设日志区利用阈值时，将由空闲的日志区转化成的数据区转化为日志区。

14. 根据权利要求13所述的硬盘控制装置，其特征在于，所述缓存器件为内存，所述硬盘控制装置还包括：

映射关系建立单元，用于建立所述数据和所述日志区空间的映射关系。

15. 根据权利要求14所述的硬盘控制装置，其特征在于，

所述映射关系建立单元，还用于建立多个目标数据和所述多个目标数据分配到的日志区空间的映射关系，其中所述目标数据属于热点数据；

所述日志管理器，还用于将所述多个目标数据的多个写操作组合为一个事务；将所述事务的所有目标数据写入日志区空间，当所述事务的其中一个目标数据的写操作执行失败时，所述事务的其他目标数据执行的写操作失败。

16. 根据权利要求15所述的硬盘控制装置，其特征在于，所述硬盘控制装置还包括：

缓存单元，用于在所述内存上缓存属于所述热点数据的数据。

17. 根据权利要求16所述的硬盘控制装置，其特征在于，

所述硬盘控制装置还包括：

链表建立单元，用于根据所述多个目标数据建立数据链表，其中，所述数据链表用于管理所述目标数据，所述数据链表管理的目标数据与所述事务的目标数据相同；

链表管理单元，用于根据所述数据链表对所述目标数据进行管理；

查找单元，用于在预设释放条件下，根据所述数据链表从先到后的建立顺序，查找所述数据链表未解除管理的数据；

内存管理单元，用于在所述内存上释放目标数据链表未解除管理的数据，且在所述内

存上保留所述目标数据链表对应的目标映射关系；

其中，所述链表管理单元，还用于：

建立第二数据链表后，当所述第二数据链表管理的第二目标数据是由预先建立的第一数据链表管理的第一目标数据修改得到时，在所述第一数据链表上解除对所述第一目标数据的管理；在与所述第一数据链表对应的第一映射关系上删除所述第一目标数据的信息。

18. 根据权利要求17所述的硬盘控制装置，其特征在于，

所述查找单元，还用于当所述内存达到第一预设水位时，根据所述数据链表从先到后的建立顺序，查找所述数据链表未解除管理的数据；

所述硬盘控制装置还包括：

读取单元，用于在所述内存达到第二预设水位时，从所述日志区读取所述目标映射关系指向的数据；

映射数据写入单元，用于将所述目标映射关系指向的数据写入所述数据区；

删除单元，用于在所述内存上删除所述目标映射关系。

19. 根据权利要求17所述的硬盘控制装置，其特征在于，

所述硬盘控制装置还包括：

回收单元，用于在预设回收条件下，执行日志区数据搬迁的步骤；

在所述执行日志区数据搬迁的步骤中，所述回收单元，包括：

回收查找模块，用于查找所述映射关系；

回收判断模块，用于根据映射关系记录的信息判断与所述映射关系对应的第一日志区上的数据是否迁移完；

回收确定模块，用于若所述第一日志区上的数据未迁移完，则根据所述映射关系记录的信息，确定第一日志区上的空间利用率；

回收执行模块，用于当所述第一日志区的空间利用率小于预设利用率阈值时，将所述第一日志区的数据迁移至第二日志区，并更新与所述被搬迁的数据对应的映射关系，其中所述第二日志区为空闲的日志区或在回收日志区时使用过的日志区；

回收模块，用于当当前日志区总的空间水位达到预设空间阈值时，则停止执行日志区数据搬迁的步骤，否则继续执行日志区数据搬迁的步骤。

20. 根据权利要求19所述的硬盘控制装置，其特征在于，

所述硬盘控制装置还包括：

事务号分配单元，用于根据所述事务的写入顺序为所述事务对应的数据链表按照递增规则分配事务号；

在所述执行日志区数据搬迁的步骤中，所述查找所述映射关系，包括：

从当前事务号最小的数据链表开始，根据所述事务号由小到大的顺序查找与所述数据链表对应的映射关系。

21. 根据权利要求16所述的硬盘控制装置，其特征在于，

所述缓存单元，还用于若所述数据是热点数据，则在所述内存上保留所述数据；或者，从所述日志区读取数据到所述缓存器件缓存。

22. 根据权利要求13至21任一项所述的硬盘控制装置，其特征在于，所述热点数据包括数据大小小于预设数据阈值的数据和/或所述热点数据包括元数据。

23. 根据权利要求13至21任一项所述的硬盘控制装置,其特征在于,
所述日志管理器,还用于为所述数据在所述日志区按照顺序分配日志区空间,将所述数据顺序追加写入所述日志区空间。

24. 根据权利要求15至21任一项所述的硬盘控制装置,其特征在于,
所述日志管理器,还用于为所述映射关系在日志区上分配日志区空间;将所述映射关系写入所述映射关系分配到的日志区空间。

硬盘数据管理方法和硬盘控制装置

技术领域

[0001] 本发明涉及数据处理领域,尤其涉及一种硬盘数据管理方法和硬盘控制装置。

背景技术

[0002] 对于普通的机械硬盘,因为其依赖机械转动硬盘和移动磁头定位读写位置,所以硬盘顺序读写是最理想的读写模型。如果硬盘空间碎片化,那么写数据的时候,无法分配到连续的空间,导致磁头抖动严重,数据传输的主要时间消耗在定位磁道和扇区上,从而留给传输数据的时间很少。因为文件的数据比较离散,那么读取的这些文件的时候,效率也较低。

[0003] 因此,大多数硬盘文件系统都在尽力避免产生大量的碎片空间,但是碎片化仍然无法避免。

[0004] 如,采用COW机制可以利用硬盘顺序写的优势。当要修改写一块数据的时候,不是直接覆盖老版本的数据,而是读取老版本的数据,修改好之后,写到一个新的位置,将写数据的数据都聚合起来,顺序写到硬盘上,释放老版本的数据。因为数据的位置变化了,需要将指向数据的上一层索引块中的指针进行修改,如此递归到最顶层。这样就会释放大量的数据,导致硬盘上产生大量的碎片。

发明内容

[0005] 本发明实施例提供了一种硬盘数据管理方法和硬盘控制装置,用于高效管理硬盘上的碎片。

[0006] 本发明第一方面提供一种硬盘数据管理方法,该方法应用于包括硬盘的硬盘控制装置,硬盘包括数据区和日志区,该方法包括:

[0007] 硬盘控制装置向缓存器件写入数据,该缓存器件例如可以是内存、flash卡、固态硬盘等与硬盘不同的存储器件,然后,硬盘控制装置判断该数据是否是热点数据,其中热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据。通过在缓存器件上对写入的数据进行判断,确定该数据的类型,以对不同的数据执行不同的处理方式。

[0008] 在向硬盘写入数据时,若该数据不是热点数据,则为数据在数据区分配数据区空间,将数据写入数据区空间;若该数据是热点数据,则为数据在日志区分配日志区空间,将数据写入日志区空间。

[0009] 本发明第一方面的硬盘数据管理方法,将待写入硬盘的数据分为热点数据和非热点数据,热点数据易于导致硬盘产生碎片,将热点数据保存在日志区上,以日志方式进行管理,即使日志区上的数据频繁修改产生硬盘碎片,也方便对这些碎片进行回收等管理,而非热点数据保存在数据区,非热点数据的释放不易导致硬盘产生碎片,数据区可以无需为硬盘碎片管理分配过多资源,从而,通过在硬盘上将不同类型的数据保存在不同的区域以不同的方式进行管理,可提高硬盘上的碎片管理效率,有效对硬盘上的碎片进行管理,规避

或减少硬盘碎片产生。

[0010] 结合第一方面,在第一种可能的实现方式中,缓存器件为内存,为该数据在日志区分配日志区空间之后,第一种可能的实现方式还包括:建立该数据和日志区空间的映射关系。即硬盘控制装置为该数据在日志区分配日志区空间后,在缓存器件上建立一映射关系,为该数据和其分配到的日志区空间的对应关系,通过该映射关系记录数据在日志区的存储情况,从而可使用该映射关系对日志区的数据和缓存器件上的数据淘汰操作进行管理。其中,在第一种可能的实现方式中,缓存器件为内存,但缓存器件还可以是其它的情形。

[0011] 结合第一方面的第一种可能的实现方式,在第二种可能的实现方式中建立数据和日志区空间的映射关系,包括:建立多个目标数据和多个目标数据分配到的日志区空间的映射关系,其中目标数据属于热点数据;

[0012] 将数据写入日志区空间,包括:将多个目标数据的多个写操作组合为一个事务;将事务的所有目标数据写入日志区空间。而当事务的其中一个目标数据的写操作执行失败时,事务的其他目标数据执行的写操作失败。多个目标数据指至少两个目标数据,相应的,多个写操作指至少两个写操作。这样,引入数据库领域事务的概念,以多个热点数据为单位对进行操作,如以多个属于同一事务的热点数据建立映射关系,和以事务的所有热点数据的写操作一起执行向日志区的写操作。这样能提高数据处理的效率。

[0013] 结合第一方面的第二种可能的实现方式,第三种可能的实现方式还包括:在内存上缓存属于热点数据的数据。将热点数据缓存在内存上,例如,在将热点数据写入日志区时,还将这些数据保留在内存上或者,在向内存写入数据前从日志区读取其上的热点数据,并缓存在内存上,这样,后续向内存写入数据时,可直接在内存上对数据进行修改,数据在内存中迁移,减少了硬盘上碎片的产生,并可根据迁移情况对日志区的数据进行整理。

[0014] 结合第一方面的第三种可能的实现方式,在第四种可能的实现方式中将事务的所有目标数据写入日志区空间之前,第四种可能的实现方式还包括:

[0015] 根据多个目标数据建立数据链表,其中,数据链表用于管理目标数据,数据链表管理的目标数据与事务的目标数据相同;然后,根据数据链表对目标数据进行管理,而,根据数据链表对目标数据进行管理,包括:建立第二数据链表后,当第二数据链表管理的第二目标数据是由预先建立的第一数据链表管理的第一目标数据修改得到时,在第一数据链表上解除对第一目标数据的管理;在与第一数据链表对应的第一映射关系上删除第一目标数据的信息。这样,在内存上,通过数据链表即可管理不同事务间数据的迁移。

[0016] 根据数据链表对内存上的数据进行管理的方式可以为:在预设释放条件下,根据数据链表从先到后的建立顺序,查找数据链表未解除管理的数据;在内存上释放目标数据链表未解除管理的数据,且在内存上保留目标数据链表对应的目标映射关系。通过释放目标数据链表上的数据可扩大内存管理数据的容量。系统通过目标映射关系的查询,即可从日志区上读取到对应的数据。

[0017] 结合第一方面的第四种可能的实现方式,在第六种可能的实现方式中在预设释放条件下,根据数据链表从先到后的建立顺序,查找数据链表未解除管理的数据,包括:当内存达到第一预设水位时,根据数据链表从先到后的建立顺序,查找数据链表未解除管理的数据;另外,该方法还包括第二阶段的内存数据淘汰,即在内存上释放目标数据链表未解除管理的数据之后,该方法还包括:在内存达到第二预设水位时,从日志区读取目标映射关系

指向的数据;将目标映射关系指向的数据写入数据区;在内存上删除目标映射关系。通过两阶段的内存数据淘汰机制可扩大内存对数据的管理容量,而在第二阶段的内存数据淘汰中,此时目标映射关系指向的数据为不活跃的数据,被修改的可能性较低,可将这些数据从日志区迁移到数据区保存,这不会过多增加数据区的碎片。

[0018] 结合第一方面的第四种可能的实现方式,在第六种可能的实现方式该方法还包括:根据事务的写入顺序为事务对应的数据链表按照递增规则分配事务号。通过为数据链表分配事务号,即可根据事务号来对数据链表进行管理,提高了管理效率。如,从当前事务号最小的数据链表开始,根据事务号由小到大的顺序查找数据链表未解除管理的数据,这样即可实现根据数据链表从先到后的建立顺序,查找数据链表未解除管理的数据。

[0019] 结合第一方面的第四种可能的实现方式,在第七种可能的实现方式中该方法还包括:在预设回收条件下,执行日志区数据搬迁的步骤。例如执行日志区数据搬迁的步骤,包括:

[0020] 查找映射关系;

[0021] 根据映射关系记录的信息判断与映射关系对应的第一日志区上的数据是否迁移完;

[0022] 若第一日志区上的数据未迁移完,则根据映射关系记录的信息,确定第一日志区上的空间利用率;

[0023] 当第一日志区的空间利用率小于预设利用率阈值时,将第一日志区的数据迁移至第二日志区,并更新与被搬迁的数据对应的映射关系,其中第二日志区为空闲的日志区或在回收日志区时使用过的日志区;

[0024] 当当前日志区总的空间水位达到预设空间阈值时,则停止执行日志区数据搬迁的步骤,否则继续执行日志区数据搬迁的步骤。

[0025] 结合第一方面的第七种可能的实现方式,在第八种可能的实现方式中该方法还包括:根据事务的写入顺序为事务对应的数据链表按照递增规则分配事务号,根据事务号来对数据链表进行管理,可提高管理效率。例如,从当前事务号最小的数据链表开始,根据事务号由小到大的顺序查找与数据链表对应的映射关系,即可实现对映射关系的查找。

[0026] 结合第一方面的第七种可能的实现方式,在第九种可能的实现方式中预设回收条件包括定时器超时、对内存数据的回收操作完成、日志区总水位达到预设水位阈值中的至少一个。

[0027] 结合第一方面的第一种可能的实现方式,在第十种可能的实现方式中,在内存上缓存属于热点数据的数据,有多种方式,例如判断数据是否是热点数据之后,若数据是热点数据,则在内存上保留该数据;或者,向缓存器件写入数据之前,从日志区读取数据到缓存器件缓存。

[0028] 结合第一方面或第一方面的第二种至第十种任一可能的实现方式,在第十一种可能的实现方式中热点数据包括数据大小小于预设数据阈值的数据和/或热点数据包括元数据。该预设数据阈值例如可以是128KB或其它的空间大小,具体的数值可根据业务类型来调整,若数据的数据大小小于该预设数据阈值,则该数据的频繁释放可能使硬盘产生大量碎片。而元数据包括对数据的管理数据,例如保存数据地址的间接块,和保存对象管理结构的元数据块。元数据也可导致硬盘产生大量碎片。对这些热点数据要筛选出,以保存在日志

区。

[0029] 结合第一方面或第一方面的第二种至第十种任一可能的实现方式,在第十二种可能的实现方式中为数据在日志区分配日志区空间,将数据写入日志区空间,包括:为数据在日志区按照顺序分配日志区空间,将数据顺序追加写入日志区空间。这样即可实现数据在日志区的顺序读写,从而在搬迁日志区的数据时,没有元数据的开销。整理开销比较小,有效保证了系统性能的稳定性的。

[0030] 结合第一方面或第一方面的第二种至第十种任一可能的实现方式,在第十三种可能的实现方式中该方法还包括:当数据区的空间利用率大于预设数据区利用阈值时,将当前空闲的日志区转化为数据区;当日志区的空间利用率大于预设日志区利用阈值时,将由空闲的日志区转化成的数据区转化为日志区。这样,日志区和数据区相互转化以适配系统容量的变化。可灵活适应具体的使用场景,提高了硬盘的使用效率。

[0031] 结合第一方面或第一方面的第二种至第十种任一可能的实现方式,在第十四种可能的实现方式中硬盘还包括超级块,每个日志区分配有标识信息,超级块用于在日志区被修改后记录被修改的日志区的标识信息。通过超级块来进一步对日志区进行管理,例如若系统断电或崩溃恢复后,硬盘控制装置即可根据该超级块记录的信息及时查找被修改的日志区。

[0032] 结合第一方面或第一方面的第二种至第十种任一可能的实现方式,在第十五种可能的实现方式中日志区和数据区在硬盘上交替设置。这样,可使得日志区和数据区的数据设置得较靠近。

[0033] 结合第一方面或第一方面的第二种至第十种任一可能的实现方式,在第十六种可能的实现方式中硬盘还包括块组,块组包括预设数量的日志区和数据区,组块的日志区和数据区连续设置,通过组块可配合调整组块内的数据区和日志区的使用,例如,根据块组的管理信息确定空闲的目标数据区后,为数据在目标数据区分配数据区空间;从而将数据写入数据区空间之后,方法还包括:根据数据和数据区空间生成目标元数据;向缓存器件写入目标元数据,硬盘控制装置判断出元数据为热点数据后,确定目标数据区所属的目标块组;确定目标块组可用的日志区;将目标元数据写入可用的日志区。这样可以将元数据在硬盘上的位置靠近元数据对应的数据在硬盘上存储的位置,方便对数据的读写。

[0034] 结合第一方面的第二种至第十种任一可能的实现方式,在第十七种可能的实现方式中该方法还包括:为映射关系在日志区上分配日志区空间,然后,将映射关系写入映射关系分配到的日志区空间。即将映射关系也保存在日志区上,从而使得映射关系在硬盘上可靠地保存。

[0035] 本发明第二方面提供一种硬盘控制装置,该硬盘控制装置包括硬盘,该硬盘包括数据区和日志区,该硬盘控制装置具有上述方法中硬盘控制装置的功能。该功能可以通过硬件实现,也可能通过硬件执行相应的软件实现。该硬件或软件包括一个或多个与上述功能相对应的模块。

[0036] 一种可能的实现方式中,该硬盘控制装置包括:

[0037] 写入单元,用于向缓存器件写入数据;

[0038] 缓存管理器,用于判断数据是否是热点数据,其中热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据;

[0039] 数据管理器,用于若数据不是热点数据,则为数据在数据区分配数据区空间,将数据写入数据区空间;

[0040] 日志管理器,用于若数据是热点数据,则为数据在日志区分配日志区空间,将数据写入日志区空间。

[0041] 另一种可能的实现方式中,该硬盘控制装置包括:

[0042] 处理器;

[0043] 该处理器执行如下动作:向缓存器件写入数据;

[0044] 该处理器执行如下动作:判断数据是否是热点数据,其中热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据;

[0045] 该处理器执行如下动作:若数据不是热点数据,则为数据在数据区分配数据区空间,将数据写入数据区空间;

[0046] 该处理器执行如下动作:若数据是热点数据,则为数据在日志区分配日志区空间,将数据写入日志区空间。

[0047] 第三方面,本申请实施例提供一种计算机存储介质,该计算机存储介质存储有程序代码,该程序代码用于指示执行上述第一方面的方法。

[0048] 从以上技术方案可以看出,本发明实施例具有以下优点:

[0049] 在包括硬盘的硬盘控制装置上,该硬盘包括数据区和日志区,向缓存器件写入数据后,硬盘控制装置判断该数据是否是热点数据,若该数据不是热点数据,则为该数据在数据区分配数据区空间,将该数据写入数据区空间;若该数据是热点数据,则为该数据在日志区分配日志区空间,将该数据写入日志区空间。这样,将待写入硬盘的数据分为热点数据和非热点数据,热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据,热点数据易于导致硬盘产生碎片,将热点数据保存在日志区上,以日志方式进行管理,即使日志区上的数据频繁修改产生硬盘碎片,也方便对这些碎片进行回收等管理,而将非热点数据保存在数据区,非热点数据的释放不易导致硬盘产生碎片,数据区可以无需为硬盘碎片管理分配过多资源,从而,通过在硬盘上将不同类型的数据保存在不同的区域以不同的方式进行管理,可提高硬盘上的碎片管理效率,日志区对硬盘碎片的高效管理,可减少硬盘碎片的产生。

附图说明

[0050] 图1为本发明实施例提供的日志区上的一个对象的逻辑视图;

[0051] 图2为本发明一实施例示出的一种硬盘数据管理方法的流程图;

[0052] 图3为图2所示实施例所涉及的数据在内存中迁移的示意图;

[0053] 图4为图2所示实施例所涉及的数据在内存中缓存的示意图;

[0054] 图5为本发明另一实施例提供的一种硬盘控制装置的结构示意图;

[0055] 图6为图5所示的硬盘控制装置的回收单元的结构示意图;

[0056] 图7为本发明另一实施例提供的一种硬盘控制装置的硬件结构示意图。

具体实施方式

[0057] 本发明实施例提供了一种硬盘数据管理方法和硬盘控制装置,于高效管理硬盘上

的碎片。

[0058] 一、本发明实施例的硬盘数据管理方法所涉及的实施环境

[0059] 本发明实施例的一种硬盘数据管理系统,该硬盘数据管理系统包括硬盘、内存,该内存可作为缓存器件,该硬盘被划分为数据区Date zone和日志区Journal zone,其中该日志区以日志方式对其上的数据进行管理。

[0060] 在硬盘数据管理系统向该硬盘写入数据前,向先作为缓存器件的内存写入该数据,若硬盘数据管理系统判断出该数据为热点数据,该热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据,这些热点数据在硬盘的数据区上修改后将导致硬盘产生大量碎片。所以,向内存写入数据后,若该数据不是热点数据,则为该数据在数据区分配数据区空间,然后,将该数据写入该数据区空间。若该数据是热点数据,则为数据在日志区分配日志区空间,将数据写入日志区空间。

[0061] 其中,数据区空间是数据区上的存储空间,可以是一个数据区上的部分空间,也可以是一个数据区上的全部空间。日志区空间是数据区上的存储空间,可以是一个日志区上的部分空间,也可以是一个日志区上的全部空间。

[0062] 这样,将待写入硬盘的数据分为热点数据和非热点数据,热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据,热点数据易于导致硬盘产生碎片,将热点数据保存在日志区上,以日志方式进行管理,即使日志区上的数据频繁修改产生硬盘碎片,也方便对这些碎片进行回收等管理,而非热点数据保存在数据区,非热点数据的释放不易导致硬盘产生碎片,数据区可以无需为硬盘碎片管理分配过多资源,从而,通过在硬盘上将不同类型的数据保存在不同的区域以不同的方式进行管理,可提高硬盘上的碎片管理效率,有效对硬盘上的碎片进行管理,通过日志区的管理也达到了规避硬盘碎片产生的效果。

[0063] 硬盘的日志区和数据区的设置可具有多种方式,如下进行详细的描述,以作为其中的一种实现方式。

[0064] 将硬盘划分为数据区和日志区两种类型的区域,对数据区和日志区的空间大小本发明实施例不作具体限定,例如可以是256M。该数据区和日志区可以是交替设置,如表一所示,表一为一种硬盘空间布局的一示例,硬盘被划分为超级块、数据区和日志区。可选地,在日志区集合中0.1%的比例作为固定的固定日志区,其均匀间隔的分布在硬盘上,这种类型的日志区只能作为日志区使用,而其它的日志区在硬盘空间不足的时候,可以转化为数据区。

[0065] 表一

[0066]

Super block 超级 块	Date zone 数据区	Journal zone 日志 区	Date zone 数据区	Journal zone 日志 区
------------------------	------------------	-------------------------	-------	-------	------------------	-------------------------

[0067] 数据区和日志区在硬盘上具有多种设置方式,上述的数据区和日志区交替设置只是其中的一种方式,本发明对此不作具体限定,例如还可以是数据区连续设置在硬盘的一

区域,日志区连续设置在硬盘的另一区域,或者多个数据区连续设置为数据区组,多个日志区连续设置为日志区组,然后数据区组和日志区组交替设置,等等。

[0068] 数据区可以保存非热点数据,例如将大于128KB的数据直接写入数据区。其中,在数据区,将数据写入数据区后,将产生管理数据的元数据,这些元数据可写入缓存器件后,再为其在日志区分配空间,以进行存储。

[0069] 日志区可以保存热点数据,例如将小于128KB的数据和元数据保存到日志区中,在有的实施例中,热点数据可以是以追加的方式保存在日志区上,在有的实施例中,还可以对日志区按照顺序分配一个标识信息ID。在日志区,以日志方式对数据进行管理。

[0070] 在有的实施例中,在日志区,按照顺序追加写方式进行io处理,当一个数据块需要写入日志区,从该日志区上一次写入的尾部分配空间,当日志区无法容纳一个数据块的时候,重新选择一个空闲最大的日志区进行追加写。

[0071] 日志区的布局如表二所示,其中Journal ctrl为标识信息ID,Map为映射关系,Record为数据,该数据例如为小于128KB的数据和元数据。

[0072] 表二

[0073]

Journal ctrl 标 识信息 ID	Map 映 射关系	Record 数据	Record 数据	Record 数据	Map 映 射关系	Record 数据	Record 数据	Record 数据
--------------------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------

[0074] 如表一所示,在有的实施例中,硬盘控制装置的硬盘上还设有超级块,在数据写入日志区后,日志区被修改,超级块将会记录被修改的日志区的标志信息ID。例如,在一批向日志区的写操作组合为一个事务时,一个事务的数据保存到硬盘之后,可将这个事务修改的日志区的标识信息ID记录到超级块对应的bitmap中。

[0075] 对于超级块的空间大小,本发明实施例不作具体限定,可根据设备具体情况进行调整,例如,一个4T的盘,日志区有 $4T/256M/2=8192$ 个,超级块中需要一个1024B来记录日志区的总体的使用情况。

[0076] 超级块的布局可以如表三所示。其中Super Blkctrk用于记录总体的管理信息,例如硬盘已经使用容量,总容量,总的空闲容量,日志区个数,数据区个数。Journalbitmap用于记录已经处理的事务号。Super blkctrk和Journalbitmap的容量可以分别是4K

[0077] 表三

[0078]

Super blkctrk	Journalbitmap
---------------	---------------

[0079] 为了让分配到数据区和日志区的数据靠近,在有的实施例中,可以将多个日志区和数据区组合为一个组块,组块上的日志区和数据区连续设置。每个块组有一个空间管理对象,采用位图文件bitmap的方式管理这个块组中的硬盘空间的使用情况。例如,可以将连续设置的16个日志区和数据区组合为一个块组。

[0080] 表四和表五示出了组块、数据区和日志区三者的关系。表四是以组块为单位对硬盘的布局的示意,表五是对表四中的组块1的布局示意。

[0081] 表四

[0082]

Super block 超级 块	member1 组块 1	member2 组块 2	memberN-1 组块 N-1	memberN 组块 N
------------------------	-----------------	-----------------	-------	-------	---------------------	-----------------

[0083] 表五

[0084]

Date zone1 数 据区 1	Journal zone2 日志区 2	Date zone15 数 据区 15	Journal zone16 日志区 16
----------------------	------------------------	-------	-------	------------------------	--------------------------

[0085] 如表二所示,日志区上还存储有映射关系Map,该该映射关系可以是用于记录日志区上的数据和该数据分配到的日志区空间的对应关系,如图1所示,其示出了硬盘对象的逻辑视图,根据该图对映射关系进行举例说明。

[0086] 如图1所示,其示出了日志区上的一个对象。一个对象可以划分为多个层次,最底层层级level为0,对应对象的数据块。level0之上是间接块,level为1。最上层是对象管理结构所在块,level为2。数据块很多的情况下,一个间接块无法保存这么多数据块的地址指针,此时需要多个间接块,对象的层数也增加了。同一层的块按照从左至右编号,例如最底层的数据块的blkid依次编号为0、1、2和3。

[0087] 事务中在修改数据块的时候,需要将数据和该数据分配到的日志区空间的关系记录到一个映射关系中。例如,一个事务创建了图1中的对象,那么需要在映射关系中记录如表六的信息。在表六中,该映射关系中每一列记录的数据的信息类型依次为objsetid,objid,levelid,blkid,journalid,offset,size。

[0088] 其中,objsetid指对象集ID,objid指对象ID,levelid指数据块所在的层数,blkid指数据块在所在层数,从左到右的序号,journalid指数据块写入的日志区的id,offset指数据块写入日志区的相对偏移,size指数据块写入的大小。

[0089] 可以理解,映射关系记录的信息类型可以是包括上述的所有信息类型,也可以包括上述信息类型的部分类型,还可以包括更多的其它的信息类型,本发明实施例对此不作具体限定。

[0090] 表六

	<objsetid, objid, levelid, blkid, journalid, offset, size>
	<1,0,0,0,1,300,4096>
	<1,0,0,1,1,400,4096>
[0091]	<1,0,0,2,1,500,4096>
	<1,0,0,3,1,600,4096>
	<1,0,1,0,1,200,4096>
	<1,0,2,0,1,100,4096>

[0092] 可以理解,在有的实施例中,该缓存器件可以由其它器件代替内存,例如Nvdim、Flash卡、SSD(固态硬盘,Solid State Drives)等。可以理解,在有的实施例中,硬盘上可以不包括日志区,而将热点数据保存在缓存器件上,本发明实施例对此不做具体限定。

[0093] 可以理解,本发明实施例的硬盘控制装置可以使用在计算机、服务器等设备上,本发明实施例对此不做具体限定。

[0094] 图2是根据一示例性实施例示出的一种硬盘数据管理方法的流程图。该方法应用于硬盘控制装置上,该硬盘控制装置包括硬盘,该硬盘包括数据区和日志区。结合上述描述的第一部分,即本发明实施例的硬盘数据管理方法所涉及的实施环境,以硬盘控制装置执行本发明实施例提供的方法的角度为例,参见图2,本发明实施例提供的方法流程包括:

[0095] 步骤201:向内存写入数据;

[0096] 在设备向硬盘写入数据前,硬盘控制装置先将该数据写入缓存器件,以进行写入硬盘前的管理,例如使用硬盘控制装置的缓存管理器对该数据进行管理。

[0097] 该缓存器件可以是内存,或者flash闪存等缓存器件,本发明实施例对此不作具体限定。

[0098] 在本发明实施例中,以缓存器件为内存进行说明。该写入内存的数据包括硬盘控制装置的所有外部数据和向硬盘的数据区写数据时产生的元数据。

[0099] 其中,向内存写入数据包括修改写和创建写两种方式,修改写即向内存写入的数据为对内存上的数据进行修改,新的数据覆盖被修改的数据,创建写即向内存写入新的数据,内存上没有缓存该新的数据的原始数据。

[0100] 步骤202:判断该数据是否是热点数据,若该数据不是热点数据,则执行步骤203,若该数据是热点数据,则执行步骤204。

[0101] 在向内存写入数据后,硬盘控制装置判断该数据是否是热点数据,如通过硬盘控

制装置的缓存管理器模块进行判断。硬盘控制装置判断待写入硬盘的数据是否是热点数据后,根据判断结果执行不同的处理方式。

[0102] 其中,热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据。例如,在本实施例中,该热点数据可以是指数据大小小于预设数据阈值的数据,和/或该热点数据也可以是指元数据。

[0103] 数据大小小于预设数据阈值的数据为小数据,该小数据频繁修改和释放容易产生硬盘碎片,而数据大小大于一定预设数据阈值的数据即使在硬盘上频繁修改也不会产生大量的硬盘碎片。其中,该预设数据阈值的设定与业务模型有关,例如可以设为64KB、128KB等。

[0104] 而元数据为记录对象管理结构和记录数据块地址的数据块。当要修改硬盘数据区上的数据块时,因数据区一般使用的是COW机制,即,当要修改写一块数据的时候,不是直接覆盖老版本的数据,而是读取老版本的数据,修改好之后,写到一个新的位置,释放老版本的数据。因为数据的位置变化了,需要将指向数据的上一层索引块中的指针进行修改,即要修改元数据,修改后的新的元数据分配到新空间,而被修改的旧的元数据需要释放掉,如此递归到最顶层。这样就会因元数据的修改,要释放大量的数据,在硬盘上释放的数据的位置就产生碎片,从而加速了硬盘碎片化的过程。

[0105] 从而,本发明实施例将数据大小小于预设数据阈值的数据和/或元数据归为热点数据,这些数据容易导致硬盘产生碎片,需要对其进行相应的管理。

[0106] 步骤203:为数据在数据区分配数据区空间,将该数据写入数据区空间。

[0107] 判断为不是热点数据的数据因其不容易使硬盘产生碎片,从而可将其保存在硬盘上的数据区。其中,数据区是硬盘上用于存储数据的区域,该区域可使用COW机制对其上的数据进行管理。数据区可以是硬盘上具有预设空间大小的区域,该预设空间大小例如可以是256M。对该数据区的描述可参考上述实施环境部分对数据区的描述。

[0108] 例如,当一数据的大小大于128KB时,缓存管理器判断该数据不是热点数据,则硬盘控制装置的数据区管理器模块可为该数据在数据区上分配数据区空间,然后将该数据写入分配到的该数据区空间。

[0109] 其中,将数据写入数据区后会产生元数据。该元数据用于记录该数据的地址,从而方便对该数据的管理。例如,组织多块数据的时候一般采用多级索引的方式,也就是在数据块的上一层分配一个索引块,其内容是记录数据块的地址,通过该索引块可以将多个数据块拼接为一个逻辑上连续的对象。这个索引块就是元数据的一种类型。将数据写入数据区产生元数据后,可将该元数据写入内存,执行上述的步骤201。缓存管理器可判断出该元数据为热点数据,从而保存到日志区和内存上。

[0110] 可以理解,在有的实施例中,硬盘控制装置的硬盘还包括块组,块组包括预设数量的日志区和数据区,块组的日志区和数据区连续设置,在包括块组的硬盘控制装置中,为数据在数据区分配数据区空间的具体执行方式是,选择合适的块组,例如空间使用较少的块组或空闲数据区较多的块组,然后为该数据在该选择的块组上分配数据区的数据区空间。在分配空间后,需要记录空间的使用情况,即需要在块组的空间管理bitmap中查找和分配空闲的数据块后,修改块组的管理结构。后续即可根据块组记录的空间使用情况向属于该块组的日志区写入该元数据,如果该块组没有空闲的日志区,需要选择邻近的日志区写入

该元数据。以使该元数据与该元数据指向的数据在硬盘上的位置靠近。

[0111] 其中,该块组可参考上述实施环境部分对块组的相应描述。

[0112] 可以理解,写入数据区上的数据除了被判断为非热点数据的数据外还可以包括内存因为空间不够而淘汰过来的数据,例如,在后续内存回收第二阶段中,将映射关系指向的数据从日志区迁移到数据区。这些被迁移的数据也可以通过数据空间管理器在数据区分配数据区空间后,写入数据区空间。

[0113] 步骤204:为数据在日志区分配日志区空间。

[0114] 硬盘上还设有日志区,在日志区上通过日志方式对日志区上的数据进行管理。

[0115] 当步骤201的数据被判断为热点数据时,并且为该数据在日志区上分配日志区空间,例如将该数据传递到日志管理器,该日志管理器为该热点数据在日志区上分配日志区空间,以使该热点数据在日志区上存储。

[0116] 本发明实施例中,为了方便地对日志区上的数据进行管理,可按照写入日志区的顺序为属于热点数据的数据在日志区上分配空间。当然,在其它的实施例中,可以不为热点数据顺序分配日志区空间,本发明实施例对此不作具体限定,

[0117] 而将属于热点数据的数据保存在日志区上,当内存上的热点数据因掉电而丢失时,可将日志区上的数据读取到内存上以供设备执行操作,并且,内存和日志区配合使用,将热点数据保存在日志区,可使得内存可管理的热点数据的数据量得到扩大。

[0118] 该日志区可以是硬盘上具有一定空间大小的区域,例如可以是256M。在硬盘上可以具有多个日志区和数据区,该日志区可以按照在硬盘上的设置顺序分配有标识信息。关于日志区具体的设置方式可参考上述实施环境部分对日志区的相应描述。

[0119] 在硬盘上日志区和数据区间的设置方式,可以具有多种方式,例如日志区和数据区交替设置在硬盘上,如上述的表一所示。当然数据区和日志区也可以以其它的方式设置,本发明实施例对此不作具体限定,具体可参考上述实施环境部分对数据区和日志区设置的相应的描述。

[0120] 步骤205:建立多个目标数据和该多个目标数据分配到的日志区空间的映射关系。

[0121] 其中目标数据属于热点数据。

[0122] 向内存写入多个数据后,经过是否是热点数据的判断后,可能在内存上得到多个属于热点数据的数据。硬盘控制装置将内存上的多个属于热点数据的目标数据确定出来,以对目标数据多个一起进行管理,提高处理效率。目标数据都在日志区上分配有日志区空间,硬盘控制装置根据这些目标数据和目标数据分配到的日志区空间建立映射关系。其中这些目标数据属于热点数据,即目标数据包括元数据和/或数据大小小于预设数据阈值的数据。

[0123] 写到日志区的所有数据都需要记录数据和日志区空间的映射关系,通过映射关系对日志区上的数据进行记录,以根据该映射关系对内存上的热点数据或日志区上的数据进行管理,例如,设备根据缓存在内存上的映射关系的索引读取保存在日志区上的对应的数据,或者根据映射关系记录的数据信息对日志区进行回收,以对日志区进行碎片管理。

[0124] 该映射关系记录的信息类型可以包括objsetid,objid,levelid,blkid,journalid,offset,size等信息,

[0125] 关于该映射关系的更多的内容可参考上述实施环境部分的相应描述。

[0126] 步骤206:在内存上缓存映射关系。

[0127] 映射关系建立好后,在内存上对其进行保存,以为后续操作准备。

[0128] 在有的实施例中,也可以将该映射关系保存到日志区,在需要保存到内存上时,再从日志区上读取该映射关系,以缓存在内存上。当然,在有的实施例中,可以将该映射关系同时保存在内存和日志区上。

[0129] 步骤207:将该多个目标数据的多个写操作组合为一个事务。

[0130] 在确定出多个目标数据后,要对该多个目标数据执行写操作以写入日志区,硬盘控制装置将这多个目标数据的多个写操作组合为一个事务,以事务为单位执行写硬盘操作。其中事务是对该多个写操作组合的称呼,并非是写操作的执行。而多个目标数据指至少两个目标数据,相应的,多个写操作指至少两个写操作。

[0131] 为了提高写入效率和保证写入可靠性,设备向日志区写入数据时,一般不是只执行一次写操作,而是在一次向日志区写入数据的操作中执行多个数据的多个写操作。这些属于同一批次的写操作的多个目标数据,在写入硬盘的过程中要么全写成功,要么全写失败。将该同一批次的多个目标数据的多个写操作组合为一个事务。

[0132] 可以理解,本发明实施例对步骤207和步骤205的执行顺序不作具体限定。即多个目标数据的写操作组合为一个事务,根据这些目标数据可建立映射关系,从而一事务与一映射关系对应。

[0133] 在有的实施例中,本发明实施例的方法还包括根据该多个目标数据建立数据链表,即根据一个事务中的一系列的元数据和小于预设阈值的数据形成一个数据链表,该数据链表用于管理目标数据。在形成映射关系时,可根据该数据链表上的数据和这些数据在日志区上分配的空间建立映射关系。

[0134] 步骤208:将事务的所有目标数据写入日志区空间。

[0135] 硬盘控制装置将事务的所有目标数据写入日志区空间,多个目标数据组合为一个事务后,当事务的其中一个目标数据的写操作执行失败时,事务的其他目标数据执行的写操作失败。只有每个目标数据的写操作都执行成功,该事务的写操作才能成功。

[0136] 映射关系建立完成后,通过日志区管理器为该映射关系分配日志区空间,属于热点数据的数据也分配了日志区空间,从而可将该映射关系和目标数据都写入两者分配到的日志区空间,以在日志区上对映射关系和目标数据进行保存。

[0137] 将映射关系保存在日志区后,系统可以将该映射关系读取到内存上,从而内存可以重新获取该映射关系,这对内存掉电后的重新工作尤其有用,当然,在有的实施例中,可以不将该映射关系保存在日志区上,从而也无需为映射关系在日志区分配空间,这也能实现减少数据区上硬盘碎片的效果,本发明实施例对此不作具体限定。

[0138] 本发明实施例中,对映射关系和事务的目标数据写入日志区空间的顺序不作具体限定。

[0139] 在有的实施例中,为数据在日志区分配日志区空间,将该数据写入日志区空间的具体方式为,为该数据在日志区按照顺序分配日志区空间,将该数据顺序追加写入该日志区空间。顺序分配空间和顺序追加写入数据即在日志区的存储空间上按照先后顺序分配空间或写入数据。

[0140] 通过为该数据在日志区按照顺序分配日志区空间,将该数据顺序追加写入该日志

区空间的方式,对日志区的数据进行管理时可以顺序读取和写入数据,提高数据管理的效率,且根据顺序来确定日志区上的数据,而映射关系记录有保存在日志区的数据和这些数据的日志区空间的对应关系,使用映射关系可代替元数据的作用,无需使用元数据来确定日志区上的数据,从而没有额外的元数据管理开销。且在日志区以顺序追加方式写数据时,可方便映射关系记录数据在日志区上的存储空间的信息。

[0141] 在本发明的实施例中,本发明实施例的方法还包括:若数据是热点数据,则在内存上缓存该数据,即将事务的所有目标数据写入日志区空间,并将事务的目标数据缓存在内存上,即判断数据是否是热点数据之后,若该数据是热点数据,则在内存上保留该数据,这样,在后续操作过程中,若向内存写入数据,该写操作若是对已经缓存在内存上的目标数据的修改写,则可直接在内存上对该数据进行修改,以执行上述的根据数据链表对目标数据进行管理的步骤。从而将该数据在内存上进行保留,以使该数据在内存上迁移,因热点数据容易导致硬盘产生碎片,将这些数据缓存在内存上而不存储在数据区,可避免在数据区上因该数据的迁移产生碎片。

[0142] 在有的实施例中,也可以不执行若数据是热点数据,则在内存上缓存该数据的步骤,而是在步骤201之前,从日志区读取数据到内存进行缓存,即向缓存器件写入数据之前,从日志区读取数据到缓存器件缓存。这样也能实现直接在内存上对该数据进行修改,本完成上述的根据数据链表对目标数据进行管理的步骤。从而将该数据在内存上进行保留,以使该数据在内存上迁移,避免在数据区上因该数据的迁移产生碎片。

[0143] 在有的实施例中,硬盘控制装置还包括超级块,在每个日志区分配有标识信息时,该超级块可用于在日志区被修改后记录被修改的日志区的标识信息。从而,当内存掉电后,可以根据超级块记录的日志区的标识信息读取对应的日志区上的数据,然后,在内存上缓存读取到的日志区上的数据,以继续执行后续数据操作。

[0144] 例如,将事务的所有目标数据和映射关系写入日志区空间后,当一个事务的所有写操作全部完成的时候,将修改的日志区的标识信息记录到超级块的bi tmap中。

[0145] 可以理解,在有的实施例中,硬盘还包括块组,该块组包括预设数量的日志区和数据区,组块的日志区和数据区连续设置,此时,如上所述,为数据在数据区分配数据区空间,具体可以是:根据块组的管理信息确定空闲的目标数据区;为数据在目标数据区分配数据区空间。

[0146] 从而,将数据写入数据区空间之后,根据数据和数据区空间会生成目标元数据,该目标元数据用于记录该数据区上的数据的地址,以对方便对该数据进行管理和查询。

[0147] 生成元数据后,向内存写入目标元数据;判断出元数据为热点数据后,确定目标数据区所属的目标块组;确定目标块组可用的日志区,即查找该目标块组上可用的日志区,然后将目标元数据写入该可用的日志区。如果没有则查找目标数据区或目标块组附近最近的可用日志区。

[0148] 其中,向日志区写入数据是以事务为单位时,为事务的多个目标数据分配日志区空间可以使用上述的方法,将目标数据的元数据分配得靠近该元数据指向的数据。

[0149] 这样,根据上述方法将元数据和该元数据指向的数据存储在硬盘上后,该元数据的存储位置靠近该元数据指向的数据的存储位置,从而减少磁头的移动距离,提高硬盘的读写效率。

[0150] 在本发明有的实施例中,对内存上的属于同一事务的多个目标数据是以链表的方式进行管理。如上所述,本发明实施例的方法在确定出多个目标数据后,可根据该多个目标数据建立数据链表,数据链表用于管理目标数据。即每一个事务的数据和元数据提交到硬盘之后,对应日志区的数据以链表方法管理起来,

[0151] 因为这些目标数据的写操作组合为一个事务,从而一个事务对应于一个数据链表,数据链表管理的目标数据与事务的目标数据相同。且根据这些目标数据建立有一映射关系,从而一个数据链表对应一映射关系,该映射关系记录了该数据链表管理的数据在日志区上的存储情况。

[0152] 根据数据链表即可对目标数据进行管理,具体的管理方法如下:

[0153] 根据上述的执行步骤,在内存上确定出属于热点数据的多个第一目标数据后,根据这些第一目标数据建立第一数据链表,这些第一目标数据属于第一事务,即这些第一目标数据在写入日志区时,是按照同一写入批次写入日志区的,只要有一个第一目标数据写入失败,则第一事务的其他数据写入失败。这些第一目标数据也建立有一第一映射关系。该第一映射关系缓存在内存上。例如,第一数据链表管理内存上的第一目标数据A1、第一目标数据B1、第一目标数据C1、第一目标数据D1。对应的第一映射关系记录有第一目标数据A1、第一目标数据B1、第一目标数据C1、第一目标数据D1与这些数据在日志区分配到的日志区空间的关系。后续过程中,硬盘控制装置根据多个第二目标数据建立第二数据链表,该多个第二目标数据属于热点数据,同时属于第二事务,根据该多个第二目标数据建立有第二映射关系,该数据链表例如可管理第二目标数据A2、第二目标数据E1、第二目标数据F1、第二目标数据D2。

[0154] 当第二数据链表管理的第二目标数据是由预先建立的第一数据链表管理的第一目标数据修改得到时,在第一数据链表上解除对第一目标数据的管理;在与第一数据链表对应的第一映射关系上删除第一目标数据的信息。这样,就实现了目标数据在数据链表上和映射关系上的迁移。例如,当第二数据链表管理的第二目标数据A2是由第一数据链表的第一目标数据A1修改得到时,因内存上的第一目标数据A1被修改为第二目标数据A2了,第一目标数据A1不再需要,从而可解除第一数据链表对第一目标数据A1的管理,相应的,第一映射关系上记录的关于第一目标数据A1的信息也可以删除。因第一映射关系记录的第一目标数据A1的信息被删除了,在日志区回收的时候,根据该第一映射关系,即可判断出与该第一目标数据A1对应的日志区有碎片产生,并在迁移合并日志区上的数据时,因从第一映射关系上读取不到第一目标数据A1的信息,从而可不将日志区上的第一目标数据A1搬迁到新的日志区,即日志区上的第一目标数据A1可释放。

[0155] 图3为数据在内存中迁移的示意图,其中,各个事务包括多个目标数据,在图3只是对其中的部分数据块进行了标号标识。如图3所示,因事务不断产生,后面的事务因为io的局部性,修改了第一事务中管理的数据。例如第二事务修改了第一目标数据A1得到第二目标数据A2,第五事务修改了第一目标数据B1得到第五目标数据B2,第四事务修改了第一目标数据C1得到第四目标数据C2,第二事务修改了第一目标数据D1得到第二目标数据D2,第三事务修改了第二目标数据D3得到第三目标数据D3。这样根据第一事务对应的和第一数据链表可知,第一事务的所有数据全部迁移到后面的事务,对应的,日志区可以全部释放第一事务的目标数据。如表7所示,在日志区1上存储有第一事务的目标数据,在第一事务的所有

数据全部迁移到后面的事务后,日志区1上面的数据块全部迁移到后面的事务,写入其它的日志区中,即其它的日志区保存有日志区1的数据对应的新版本的数据,这个时候,在日志区回收时日志区1就可以释放出来,作为空的日志区。其它的目标数据的迁移类似,执行到第五事务的时候,内存中实际的缓存情况如图4所示。

[0156] 表7

[0157]	Journal zone1	MAP	A1	B1	C1	D1
	Journal zone2	MAP	A2	*	*	D2
	Journal zone3	MAP	*	*	*	D3
	Journal zone4	MAP	*	C2	*	*
	Journal zone5	MAP	B2	*	*	*

[0158] 为了更好地根据数据链表对目标数据进行管理,在有的实施例中,本发明实施例的方法中,还包括根据事务的写入顺序为事务对应的数据链表按照递增规则分配事务号,事务的写入顺序是指不同事务间写入日志区的先后顺序,先写入的事务对应的数据链表分配到的事务号较小,跟着写入日志区的事务对应的数据链表的事务号增加一个单位,这样每个数据链表即可有相应的标识,这些标识还具有递增的规律。从而,根据事务号可方便地确定出内存上的数据链表的建立先后顺序。

[0159] 例如,首先建立第一事务的数据链表,该第一事务先写入日志区,从而为该第一事务的数据链表分配事务号1,然后根据多个第二目标数据建立第二事务的数据链表,该第二事务的数据将在第一事务之后写入日志区,从而为该第二事务对应的第二数据链表分配事务号2,类似地,为第三数据链表分配事务号3,如此类推。

[0160] 上述即为根据数据链表对数据进行管理的部分内容。属于热点数据的数据写入日志区,通过在内存上缓存该属于热点数据的数据或者从日志区读取数据后,将该读取的数据缓存在内存上,从而可在内存上根据数据链表对数据进行管理,让热点数据在内存上迁移,减少该数据在硬盘上的迁移,以减少该数据使硬盘产生的碎片。

[0161] 将数据缓存在内存上,还可以在读取数据时,直接从内存读取数据,避免读取日志区中的数据。

[0162] 并且通过在内存上根据数据链表对数据进行管理,以对映射关系进行建立和改变,根据映射关系即可对日志区的数据和碎片进行整理,以避免或者减少硬盘的碎片,还能提高碎片处理的效率。为了充分利用内存上的空间,可以对缓存在内存上的热点数据进行

释放,回收内存的空间,以使得内存缓存其它更多的数据。所以,本发明实施例的方法,在预设释放条件下,可对内存空间进行回收。例如,达到系统内存的回收水位时,触发缓存淘汰。下面示出一示例性的回收方法,该回收方法分为两阶段。

[0163] 第一阶段

[0164] 当内存达到第一预设水位时,从当前事务号最小的数据链表开始,根据事务号由小到大的顺序查找数据链表未解除管理的数据。然后,对查找到的目标数据链表,在内存上释放目标数据链表未解除管理的数据,且在内存上保留目标数据链表对应的目标映射关系。

[0165] 因通过数据链表的方式对数据进行了管理,先建立的数据链表管理的数据可能因被后写入的数据的修改而在内存上淘汰掉,这些淘汰掉的数据从数据链表上已经完成了解除。对此,可参考上述对数据链表的管理方法的描述。对内存进行回收,即,查找数据链表没淘汰的数据,在内存上释放这些还没淘汰的数据。这些在内存回收时被释放的数据所属的数据链表可称为目标数据链表。而目标数据链表对应的映射关系仍然保留在内存上。第一阶段的内存回收直到内存空间达到预设的停止水位时,才停止。

[0166] 经过上述的步骤,小于等于预设数据阈值的数据和写数据产生的元数据写到日志区,同时在内存中缓存。对这些日志区和内存上的热点数据,如上,采用事务先后的顺序管理起来,这样当系统运行一段时间,缓存达到第一预设水位的时候,触发后台缓存回收线程将按照事务号从小到大的顺序回收内存空间。例如,从当前事务号最小的数据链表开始,根据事务号由小到大的顺序查找数据链表未解除管理的数据,实现了从最先建立的数据链表开始释放内存上的数据。

[0167] 越早建立的数据链表剩下的没在内存淘汰的数据是较不活跃的数据,设备对其读取修改的可能性较小,从而可将这些数据从内存释放,这样对设备的数据读取效率影响较小。而因目标映射关系保存在内存上,设备要读取数据时,若从内存上读取不上,则根据内存上保留的映射关系进行查询,若根据目标映射关系确定出要查询的数据,则可根据目标映射关系从日志区上读取对应的数据。这样,内存即可管理更多的热点数据。

[0168] 可以理解,内存达到第一预设水位,只是预设释放条件的一种,查找数据链表未解除管理的数据也可以在其它的预设释放条件下触发,例如设定的计时器到时等,本发明实施例对此不作具体限定。

[0169] 第二阶段

[0170] 内存进行第一阶段的回收后,在内存达到第二预设水位时,从日志区读取目标映射关系指向的数据;将目标映射关系指向的数据写入数据区;在内存上删除目标映射关系。

[0171] 为了进一步充分利用内存的空间,可在第二阶段对内存空间进行第二次回收。其中,第二预设水位时触发内存执行第二阶段回收的水位。第二预设水位可以与第一预设水位相同,也可以与第一预设水位不相同,本发明实施例对此不做具体限定,对第一预设水位和第二预设水位的具体数值本发明实施例也不做具体限定,例如,可根据实际内存容量和业务类型灵活设定。

[0172] 从日志区读取目标映射关系指向的数据后,可将目标映射关系指向的数据写入数据区。当内存达到第二预设水位时,此时内存运行了一定的时间,缓存在内存中的映射关系越来越多,最终也达到缓存水位的时候,需要读取这些映射关系指向的数据,写入到数据区

中。

[0173] 在内存达到第二预设水位时,因执行了上述的数据链表的管理方法,当前越早建立的数据链表的没被解除的数据被修改的可能性越小,因为,若数据链表上的管理的数据被后面写入的数据修改,则该数据从该数据链表迁移到后面事务的对应的数据链表。从而,越早建立的数据链表剩下的数据可判定为不活跃的数据,这些不活跃的数据因被修改的可能性较小,从而它们因修改产生的硬盘碎片也较少,可将目标映射关系指向的数据写入数据区。例如,根据目标映射关系从日志区读取这些数据,然后由数据区管理器在数据区上分配好数据区空间,再将这些数据写入分配到的数据区上。

[0174] 经过上述步骤的执行后,当设备要读取硬盘上的数据时,例如有对已经写入日志区的数据进行访问,硬盘控制装置可以先在内存上查找,如果在内存中命中,则可以直接返回该数据。因内存回收的第一阶段后,从内存中删除了部分热点数据,而保留了对应的映射关系,从而如果在内存中没有查到要访问的数据,但在内存缓存的映射关系中可以查找到,则可以在该映射关系对应的日志区中直接读取这些要访问的数据。如果在映射关系中,也查找不到要查找的数据,则到数据区中查找和读取数据。

[0175] 可以理解,在按照顺序追加写的方式对日志区进行写数据的实施例中,当连续多个事务的数据链表淘汰完毕之后,例如执行上述的内存第二阶段的回收后,与被删除的映射关系对应的日志区也完全释放了,可以重新作为一个空的日志区来使用。具体的释放日志区以及回收日志区的方法下面将叙及。

[0176] 在有的实施例中,执行上述方法后,本发明实施例还包括对日志区回收的操作。从而减少日志区上的碎片,充分利用日志区的空间。

[0177] 日志区的回收,可基于上述的方法,下面以按照顺序追加写的方式对日志区进行写数据的实施例对日志区的回收进行说明。即步骤208为,将事务的目标数据顺序追加写入日志区空间,即日志区上的数据是根据事务的先后顺序存储的,

[0178] 日志区的具体回收方法,例如可包括下述步骤:

[0179] A1:在预设回收条件下,执行日志区数据搬迁的步骤;

[0180] 其中,预设回收条件包括定时器超时、对内存数据的回收操作完成、日志区总水位达到预设水位阈值中的至少一个。对内存数据的回收操作完成指上述的内存回收的时候,每个阶段执行完成,都触发日志区数据搬迁步骤,即触发日志回收线程执行日志区回收流程。

[0181] A2:当当前日志区总的空间水位达到预设空间阈值时,则停止执行日志区数据搬迁的步骤,否则继续执行日志区数据搬迁的步骤。

[0182] 其中,执行日志区数据搬迁的步骤,包括:

[0183] B1:从当前事务号最小的数据链表开始,根据事务号由小到大的顺序查找与数据链表对应的映射关系。

[0184] 在上述的方法执行了根据事务的写入顺序为事务对应的数据链表按照递增规则分配事务号的步骤之后,硬盘控制装置按照事务号从小到大的顺序查找与数据链表对应的映射关系,因映射关系记录了数据和该数据在日志区上分配到的空间的关系,从而可分析映射关系中的数据块在日志区上的分布情况,即可知道对应的日志区上还有多少数据没迁移,没迁移的数据位于对应的日志区上。

[0185] 而因为事务在推进的时候,使用完毕一个日志区之后才切换到下一个日志区,因此连续事务的数据会写到连续的日志区中。根据事务号从小到大的顺序分析映射关系,即可获得到连续的日志区上的数据存储情况。

[0186] B2:根据映射关系记录的信息判断与该映射关系对应的第一日志区上的数据是否迁移完毕;

[0187] 在根据数据链表对内存上的数据进行管理时,对在内存上迁移的数据,也要修改对应的映射关系,若数据链表记录的信息反映某日志区上的数据以在内存上迁移完毕,例如该日志区对应的数据都在其它的日志区有新的版本,或者内存回收的第二阶段后,已经将对应的日志区上的数据搬迁到数据区,此时,判断出该映射关系对应的日志区的数据已经迁移完。

[0188] B3:若第一日志区的数据迁移完毕,则回收该第一日志区;

[0189] 若第一日志区的数据已经搬迁完毕,则回收该第一日志区。在通过超级块记录日志区的使用情况的实施例中,此时可将超级块中的与第一日志区对应的信息清楚掉。

[0190] B4:若第一日志区上的数据未迁移完,则根据映射关系记录的信息,确定第一日志区上的空间利用率;

[0191] 因映射关系记录了数据和该数据在日志区上分配到的空间的关系,从而可根据映射关系记录的信息分析出第一日志区上的空间利用率。

[0192] B5:当第一日志区的空间利用率小于预设利用率阈值时,将第一日志区的数据迁移至第二日志区,并更新与被搬迁的数据对应的映射关系。

[0193] 其中第二日志区为空闲的日志区或在回收日志区时使用过的日志区,例如前次日志区回收时使用过的日志区。该预设利用阈值可以根据具体情况设定,例如,在日志区使用较多时,但是每个日志区的利用率较低的时候可以降低该预设利用阈值。例如该预设利用阈值可设为50%等,本发明实施例对此不作具体限定。更新与被搬迁的数据对应的映射关系,可以是数据在日志区上搬迁后,在与该数据相关的映射关系上及时更新该数据和新的日志区空间的对应关系。

[0194] 通过上述的日志区回收方法,即可实现了日志区的回收,减少日志区上的碎片,充分利用了日志区的空间。而因日志区是顺序读写的,与数据链表的事务号从小到大的顺序对应,从而在可根据数据链表的事务号大小查询对应的映射关系,并搬迁数据,此时无需元数据来记录日志区的数据的位置,从而减少了元数据的开销。

[0195] 本发明实施例的方法,在包括硬盘的硬盘控制装置上,该硬盘包括数据区和日志区,向缓存器件写入数据后,硬盘控制装置判断该数据是否是热点数据,若该数据不是热点数据,则为该数据在数据区分配数据区空间,将该数据写入数据区空间;若该数据是热点数据,则为该数据在日志区分配日志区空间,将该数据写入日志区空间。这样,将待写入硬盘的数据分为热点数据和非热点数据,热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据,热点数据易于导致硬盘产生碎片,将热点数据保存在日志区上,以日志方式进行管理,即使日志区上的数据频繁修改产生硬盘碎片,也方便对这些碎片进行回收等管理,而将非热点数据保存在数据区,非热点数据的释放不易导致硬盘产生碎片,数据区可以无需为硬盘碎片管理分配过多资源,从而,通过在硬盘上将不同类型的数据保存在不同的区域以不同的方式进行管理,可提高硬盘上的碎片管理效率,日志区

对硬盘碎片的高效管理,可减少硬盘碎片的产生。

[0196] 在日志区中的数据因为热点数据的局部性,被迁移走了,实际在整理日志区的时候,读取的数据较少,从而日志区的数据管理效率较高。而将热点数据保存在日志区,因热点数据容易产生碎片,从而碎片集中在日志区,进一步提高了对碎片整理的效率。

[0197] 在以顺序追加写的方式向日志区写入数据时,在日志区只需要顺序读取数据和顺序写入数据,搬迁数据没有产生元数据的开销,日志区进行碎片整理效率高。通过整理日志区的数据,可有效避免或者减少硬盘碎片。

[0198] 在有的实施例中,也可以不以顺序追加方式向日志区写入数据,在知道日志区的空间分配情况时,可以使用位图文件bitmap对应journal zone来管理。

[0199] 例如,每个bit对应一个固定的块大小,例如是4K,那么256M的journal zone需要8K来管理,每次写journal zone都需要修改这个bitmap。

[0200] 可以理解,在本发明的实施例中,数据区和日志区可以为分级存储,即数据区和日志区在不同的层级,在日志区回收的时候,将热点数据迁移到较低层级的存储层。

[0201] 可以理解,在有的实施例中,可以以其它方式查找映射关系,例如随机查询映射关系,然后根据查到的映射关系来分析对应的日志区的空间存储情况,然后进行日志区数据的搬迁,此时可不用将事务的目标数据以顺序追加的方式写入日志区空间,对具体的写入方式不做限定。但是,这样的方式可能因映射关系反应不出日志区的全部空间存储情况,而导致日志区的回收效果不够理想。

[0202] 可以理解,在本发明的包括多个日志区和多个数据区的实施例中,为了更充分地使用数据区和日志区,从而充分利用硬盘空间,本发明实施例的方法还包括数据区和日志区的转化步骤,例如,当数据区的空间利用率大于预设数据区利用阈值时,将当前空闲的日志区转化为数据区;当日志区的空间利用率大于预设日志区利用阈值时,将由空闲的日志区转化成的数据区转化为日志区。

[0203] 例如,硬盘上初始化的时候,预先设定,有一半的硬盘空间是日志区,其他的是数据区。当系统运行一段时间之后,数据区的利用率较高,按照日志区的标识信息递增顺序查找空闲的日志区,将其转化为数据区。转化为数据区之后,在包括块组的实施例中,使用块组的空间管理对象将其管理起来。将日志区的状态设置为数据区,并记录到块组的管理结构中,对此写盘保存。当数据区上的数据被删除,硬盘空间释放出来,一个由日志区转化成的数据区的空间全部释放出来,这个时候将该数据区的状态切换为日志区,并记录在块组的管理结构中,写盘保存。

[0204] 图5为根据一示例性实施例示出的一种硬盘控制装置的结构示意图,该硬盘控制装置包括硬盘,硬盘包括数据区和日志区,该硬盘控制装置用于执行上述图2对应的实施例中硬盘控制装置执行的功能。参见图5,该硬盘控制装置包括:

[0205] 写入单元501,用于向缓存器件写入数据;

[0206] 缓存管理器502,用于判断数据是否是热点数据,其中热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据;

[0207] 数据管理器503,用于若数据不是热点数据,则为数据在数据区分配数据区空间,将数据写入数据区空间;

[0208] 日志管理器504,用于若数据是热点数据,则为数据在日志区分配日志区空间,将

数据写入日志区空间。

[0209] 可选地,缓存器件为内存,硬盘控制装置还包括:

[0210] 映射关系建立单元505,用于在内存建立数据和日志区空间的映射关系;

[0211] 可选地,硬盘控制装置还包括:

[0212] 缓存单元506,用于在内存上缓存属于热点数据的数据。

[0213] 可选地,

[0214] 映射关系建立单元505,还用于建立多个目标数据和多个目标数据分配到的日志区空间的映射关系,其中目标数据属于热点数据;

[0215] 日志管理器504,还用于将多个目标数据的多个写操作组合为一个事务将事务的所有目标数据写入日志区空间,当事务的其中一个目标数据的写操作执行失败时,事务的其他目标数据执行的写操作失败。

[0216] 可选地,

[0217] 硬盘控制装置还包括:

[0218] 链表建立单元509,用于根据多个目标数据建立数据链表,其中,数据链表用于管理目标数据,数据链表管理的目标数据与事务的目标数据相同;

[0219] 链表管理单元510,用于根据数据链表对目标数据进行管理;

[0220] 查找单元511,用于在预设释放条件下,根据数据链表从先到后的建立顺序,查找数据链表未解除管理的数据;

[0221] 内存管理单元512,用于在内存上释放目标数据链表未解除管理的数据,且在内存上保留目标数据链表对应的目标映射关系;

[0222] 其中,链表管理单元510,还用于:

[0223] 建立第二数据链表后,当第二数据链表管理的第二目标数据是由预先建立的第一数据链表管理的第一目标数据修改得到时,在第一数据链表上解除对第一目标数据的管理;在与第一数据链表对应的第一映射关系上删除第一目标数据的信息

[0224] 可选地,

[0225] 查找单元511,还用于当内存达到第一预设水位时,根据数据链表从先到后的建立顺序,查找数据链表未解除管理的数据;

[0226] 硬盘控制装置还包括:

[0227] 读取单元523,用于在内存达到第二预设水位时,从日志区读取目标映射关系指向的数据;

[0228] 映射数据写入单元513,用于将目标映射关系指向的数据写入数据区;

[0229] 删除单元514,用于在内存上删除目标映射关系。

[0230] 可选地,

[0231] 硬盘控制装置还包括:

[0232] 事务号分配单元515,用于根据事务的写入顺序为事务对应的数据链表按照递增规则分配事务号;

[0233] 查找单元511,还用于从当前事务号最小的数据链表开始,根据事务号由小到大的顺序查找数据链表未解除管理的数据。

[0234] 可选地,

- [0235] 硬盘控制装置还包括：
- [0236] 回收单元516,用于在预设回收条件下,执行日志区数据搬迁的步骤；
- [0237] 如图6所示,在执行日志区数据搬迁的步骤中,回收单元516,包括：
- [0238] 回收查找模块517,用于查找映射关系；
- [0239] 回收判断模块518,用于根据映射关系记录的信息判断与映射关系对应的第一日志区上的数据是否迁移完；
- [0240] 回收确定模块519,用于若第一日志区上的数据未迁移完,则根据映射关系记录的信息,确定第一日志区上的空间利用率；
- [0241] 回收执行模块520,用于当第一日志区的空间利用率小于预设利用率阈值时,将第一日志区的数据迁移至第二日志区,并更新与被搬迁的数据对应的映射关系,其中第二日志区为空闲的日志区或在回收日志区时使用过的日志区；
- [0242] 回收模块521,用于当当前日志区总的空间水位达到预设空间阈值时,则停止执行日志区数据搬迁的步骤,否则继续执行日志区数据搬迁的步骤。
- [0243] 可选地,
- [0244] 硬盘控制装置还包括：
- [0245] 事务号分配单元515,用于根据事务的写入顺序为事务对应的数据链表按照递增规则分配事务号；
- [0246] 回收查找模块517,还用于从当前事务号最小的数据链表开始,根据事务号由小到大的顺序查找与数据链表对应的映射关系；
- [0247] 可选地,预设回收条件包括定时器超时、对内存数据的回收操作完成、日志区总水位达到预设水位阈值中的至少一个。
- [0248] 可选地,
- [0249] 缓存单元506,还用于若数据是热点数据,则在内存上缓存数据;或者,从日志区读取数据到缓存器件缓存。
- [0250] 可选地,热点数据包括数据大小小于预设数据阈值的数据和/或热点数据包括元数据。
- [0251] 可选地,
- [0252] 日志管理器504,还用于为数据在日志区按照顺序分配日志区空间,将数据顺序追加写入日志区空间。
- [0253] 可选地,
- [0254] 硬盘控制装置还包括：
- [0255] 数据区转化单元522,用于当数据区的空间利用率大于预设数据区利用阈值时,将当前空闲的日志区转化为数据区；
- [0256] 日志区转化单元524,用于当日志区的空间利用率大于预设日志区利用阈值时,将由空闲的日志区转化成的数据区转化为日志区。
- [0257] 可选地,
- [0258] 硬盘还包括超级块,每个日志区分配有标识信息,超级块用于在日志区被修改后记录被修改的日志区的标识信息。
- [0259] 可选地,日志区和数据区在硬盘上交替设置。

[0260] 可选地，

[0261] 硬盘还包括块组，块组包括预设数量的日志区和数据区，组块的日志区和数据区连续设置，

[0262] 数据管理器503，包括：

[0263] 空闲区确定模块525，用于根据块组的管理信息确定空闲的目标数据区；

[0264] 分配模块508，用于为数据在目标数据区分配数据区空间；

[0265] 硬盘控制装置还包括：

[0266] 元数据生成模块507，用于根据数据和数据区空间生成目标元数据；

[0267] 写入单元501，还用于向缓存器件写入目标元数据；

[0268] 缓存管理器判断出元数据为热点数据后，日志管理器504，还用于确定目标数据区所属的目标块组；确定目标块组可用的日志区；将目标元数据写入可用的日志区。

[0269] 可选地，

[0270] 日志管理器504，还用于为映射关系在日志区上分配日志区空间；将映射关系写入映射关系分配到的日志区空间。

[0271] 综上所述，在包括硬盘的硬盘控制装置上，该硬盘包括数据区和日志区，写入单元501向缓存器件写入数据后，缓存管理器502判断该数据是否是热点数据，若该数据不是热点数据，则数据管理器503为该数据在数据区分配数据区空间，将该数据写入数据区空间；若该数据是热点数据，则日志管理器504为该数据在日志区分配日志区空间，将该数据写入日志区空间。这样，将待写入硬盘的数据分为热点数据和非热点数据，热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据，热点数据易于导致硬盘产生碎片，将热点数据保存在日志区上，以日志方式进行管理，即使日志区上的数据频繁修改产生硬盘碎片，也方便对这些碎片进行回收等管理，而将非热点数据保存在数据区，非热点数据的释放不易导致硬盘产生碎片，数据区可以无需为硬盘碎片管理分配过多资源，从而，通过在硬盘上将不同类型的数据保存在不同的区域以不同的方式进行管理，可提高硬盘上的碎片管理效率，日志区对硬盘碎片的高效管理，可减少硬盘碎片的产生。

[0272] 图7为本发明另一实施例提供的一种硬盘控制装置的硬件结构示意图，该硬盘控制装置包括处理器CPU701、缓存器件703和硬盘702，以及硬盘控制器705和总线704。硬盘702包括数据区和日志区，在有的实施例中缓存器件例如可以是内存。

[0273] 上述实施例中由硬盘控制装置所执行的步骤可以基于该图7所示的硬盘控制装置的结构。

[0274] 该处理器701执行程序，使得硬盘控制装置执行上述硬盘数据管理方法的方法，举例各种可选设计具体如下。

[0275] 该处理器701执行程序，使得硬盘控制装置具有如下功能：向缓存器件写入数据；判断数据是否是热点数据，其中热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据；若数据不是热点数据，则为数据在数据区分配数据区空间，将数据写入数据区空间；若数据是热点数据，则为数据在日志区分配日志区空间，将数据写入日志区空间。

[0276] 一种可选设计，缓存器件为内存，该处理器701执行程序，使得硬盘控制装置具有如下功能：为数据在日志区分配日志区空间之后，建立数据和日志区空间的映射关系。

[0277] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:建立多个目标数据和多个目标数据分配到的日志区空间的映射关系,其中目标数据属于热点数据;将多个目标数据的多个写操作组合为一个事务;将事务的所有目标数据写入日志区空间,当事务的其中一个目标数据的写操作执行失败时,事务的其他目标数据执行的写操作失败。

[0278] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:在内存上缓存属于热点数据的数据。

[0279] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:将事务的所有目标数据写入日志区空间之前,根据多个目标数据建立数据链表,其中,数据链表用于管理目标数据,数据链表管理的目标数据与事务的目标数据相同;根据数据链表对目标数据进行管理;在预设释放条件下,根据数据链表从先到后的建立顺序,查找数据链表未解除管理的数据;在内存上释放目标数据链表未解除管理的数据,且在内存上保留目标数据链表对应的目标映射关系;其中,根据数据链表对目标数据进行管理,包括:建立第二数据链表后,当第二数据链表管理的第二目标数据是由预先建立的第一数据链表管理的第一目标数据修改得到时,在第一数据链表上解除对第一目标数据的管理;在与第一数据链表对应的第一映射关系上删除第一目标数据的信息

[0280] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:当内存达到第一预设水位时,根据数据链表从先到后的建立顺序,查找数据链表未解除管理的数据;

[0281] 在内存上释放目标数据链表未解除管理的数据之后,在内存达到第二预设水位时,从日志区读取目标映射关系指向的数据;

[0282] 将目标映射关系指向的数据写入数据区;

[0283] 在内存上删除目标映射关系。

[0284] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:

[0285] 根据事务的写入顺序为事务对应的数据链表按照递增规则分配事务号;

[0286] 从当前事务号最小的数据链表开始,根据事务号由小到大的顺序查找数据链表未解除管理的数据。

[0287] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:在预设回收条件下,执行日志区数据搬迁的步骤;

[0288] 执行日志区数据搬迁的步骤,包括:

[0289] 查找映射关系;

[0290] 根据映射关系记录的信息判断与映射关系对应的第一日志区上的数据是否迁移完;

[0291] 若第一日志区上的数据未迁移完,则根据映射关系记录的信息,确定第一日志区上的空间利用率;

[0292] 当第一日志区的空间利用率小于预设利用率阈值时,将第一日志区的数据迁移至第二日志区,并更新与被搬迁的数据对应的映射关系,其中第二日志区为空闲的日志区或在回收日志区时使用过的日志区;

[0293] 当当前日志区总的空间水位达到预设空间阈值时,则停止执行日志区数据搬迁的

步骤,否则继续执行日志区数据搬迁的步骤。

[0294] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:

[0295] 根据事务的写入顺序为事务对应的数据链表按照递增规则分配事务号;从当前事务号最小的数据链表开始,根据事务号由小到大的顺序查找与数据链表对应的映射关系;

[0296] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:

[0297] 预设回收条件包括定时器超时、对内存数据的回收操作完成、日志区总水位达到预设水位阈值中的至少一个。

[0298] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:

[0299] 判断数据是否是热点数据之后,若数据是热点数据,则在内存上缓存数据;或者,向缓存器件写入数据之前,从日志区读取数据到缓存器件缓存。

[0301] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:

[0302] 热点数据包括数据大小小于预设数据阈值的数据和/或热点数据包括元数据。

[0303] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:

[0304] 为数据在日志区按照顺序分配日志区空间,将数据顺序追加写入日志区空间。

[0305] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:

[0306] 当数据区的空间利用率大于预设数据区利用阈值时,将当前空闲的日志区转化为数据区;

[0307] 当日志区的空间利用率大于预设日志区利用阈值时,将由空闲的日志区转化成的数据区转化为日志区。

[0308] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:

[0309] 硬盘还包括超级块,每个日志区分配有标识信息,超级块用于在日志区被修改后记录被修改的日志区的标识信息。

[0310] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:

[0311] 日志区和数据区在硬盘上交替设置。

[0312] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:硬盘还包括块组,块组包括预设数量的日志区和数据区,组块的日志区和数据区连续设置,根据块组的管理信息确定空闲的目标数据区;为数据在目标数据区分配数据区空间;

[0313] 将数据写入数据区空间之后,根据数据和数据区空间生成目标元数据;向缓存器件写入目标元数据;判断出元数据为热点数据后,确定目标数据区所属的目标块组;确定目标块组可用的日志区;将目标元数据写入可用的日志区。

[0314] 一种可选设计,该处理器701执行程序,使得硬盘控制装置具有如下功能:为映射关系在日志区上分配日志区空间;将映射关系写入映射关系分配到的日志区空间。

[0315] 综上所述,在包括硬盘的硬盘控制装置上,该硬盘包括数据区和日志区,该处理器701向缓存器件写入数据后,该处理器701判断该数据是否是热点数据,若该数据不是热点数据,则该处理器701为该数据在数据区分配数据区空间,将该数据写入数据区空间;若该数据是热点数据,则该处理器701为该数据在日志区分配日志区空间,将该数据写入日志区空间。这样,将待写入硬盘的数据分为热点数据和非热点数据,热点数据为存储在硬盘上后在预设次数的修改和释放后能使硬盘产生预设数量碎片的数据,热点数据易于导致硬盘产生碎片,将热点数据保存在日志区上,以日志方式进行管理,即使日志区上的数据频繁修改产

生硬盘碎片,也方便对这些碎片进行回收等管理,而将非热点数据保存在数据区,非热点数据的释放不易导致硬盘产生碎片,数据区可以无需为硬盘碎片管理分配过多资源,从而,通过在硬盘上将不同类型的数据保存在不同的区域以不同的方式进行管理,可提高硬盘上的碎片管理效率,日志区对硬盘碎片的高效管理,可减少硬盘碎片的产生。

[0316] 所属领域的技术人员可以清楚地了解到,为描述的方便和简洁,上述描述的系统,装置和单元的具体工作过程,可以参考前述方法实施例中的对应过程,在此不再赘述。

[0317] 在本申请所提供的几个实施例中,应该理解到,所揭露的系统,装置和方法,可以通过其它的方式实现。例如,以上所描述的装置实施例仅仅是示意性的,例如,所述单元的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,例如多个单元或组件可以结合或者可以集成到另一个系统,或一些特征可以忽略,或不执行。另一点,所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口,装置或单元的间接耦合或通信连接,可以是电性,机械或其它的形式。

[0318] 所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

[0319] 另外,在本发明各个实施例中的各功能单元可以集成在一个处理单元中,也可以是各个单元单独物理存在,也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现,也可以采用软件功能单元的形式实现。

[0320] 所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用,可以存储在一个计算机可读取存储介质中。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质中,包括若干指令用以使得一台计算机设备(可以是个人计算机,服务器,或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括:U盘、移动硬盘、只读存储器(ROM,Read-Only Memory)、随机存取存储器(RAM,Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0321] 以上所述,以上实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

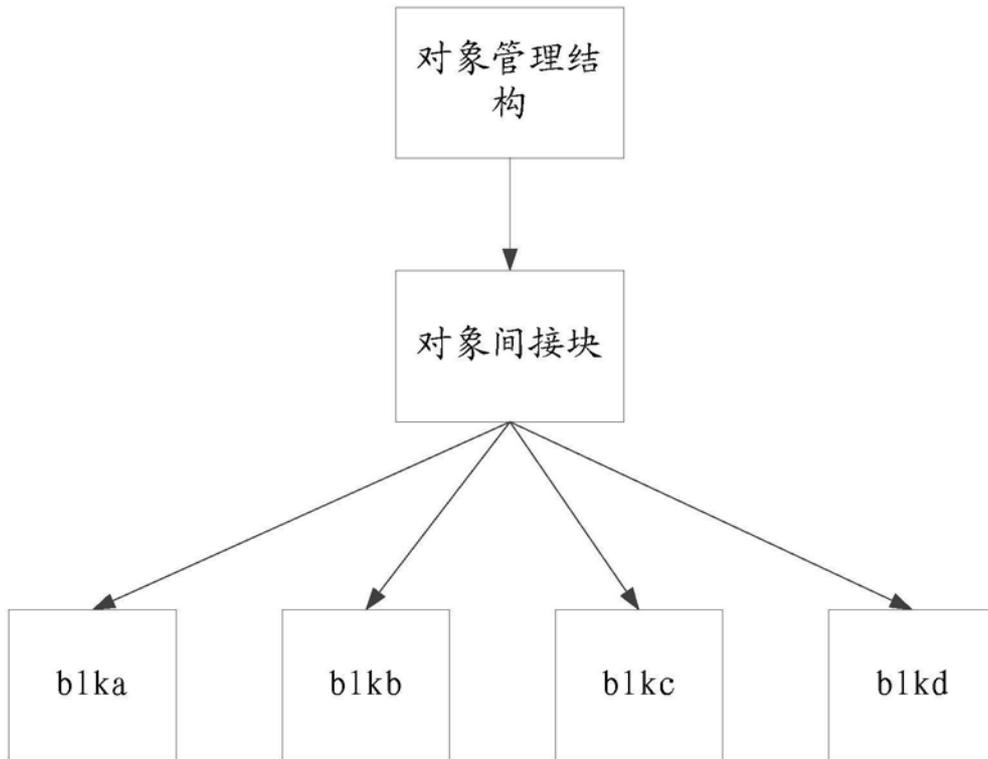


图1

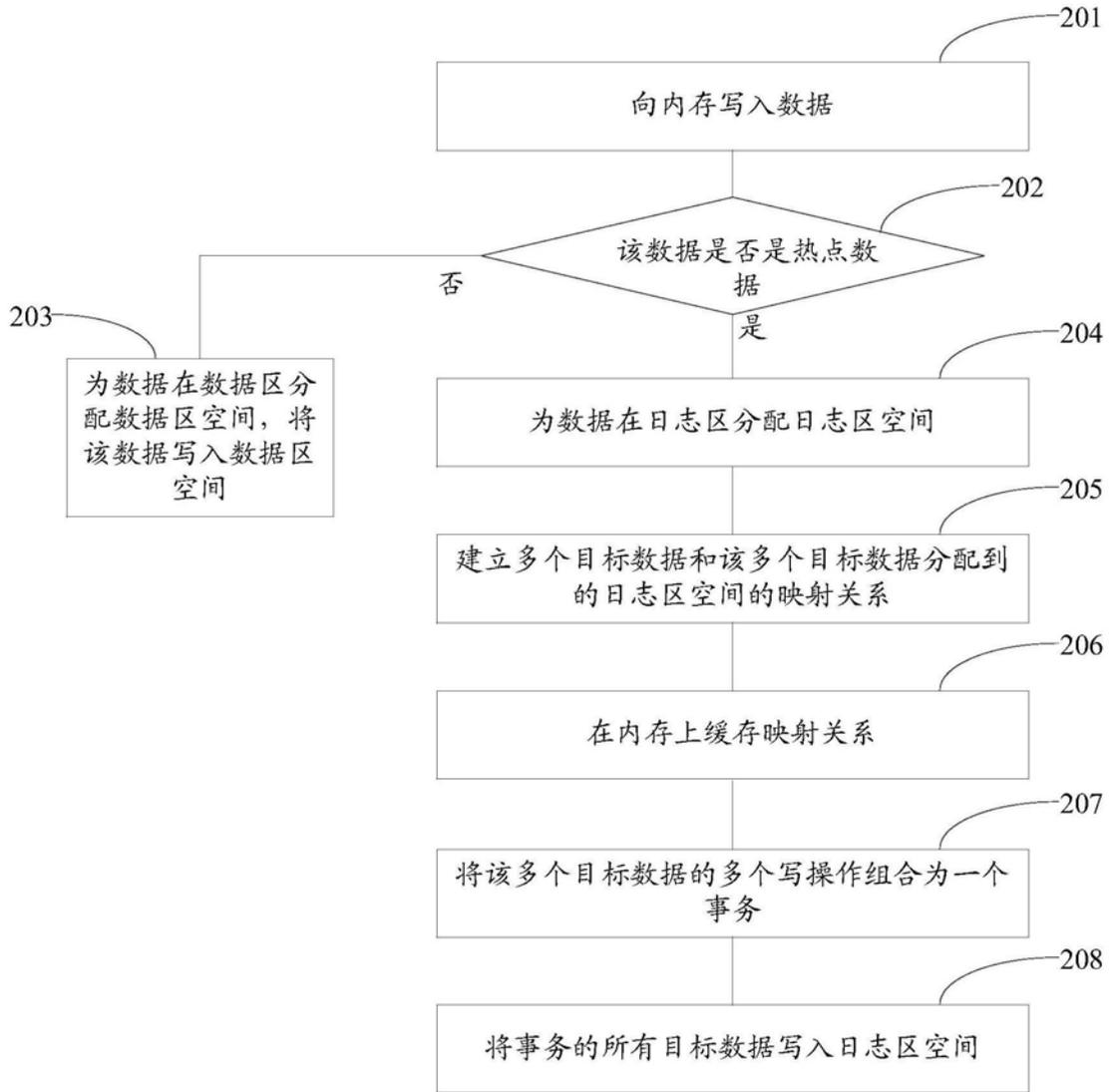


图2

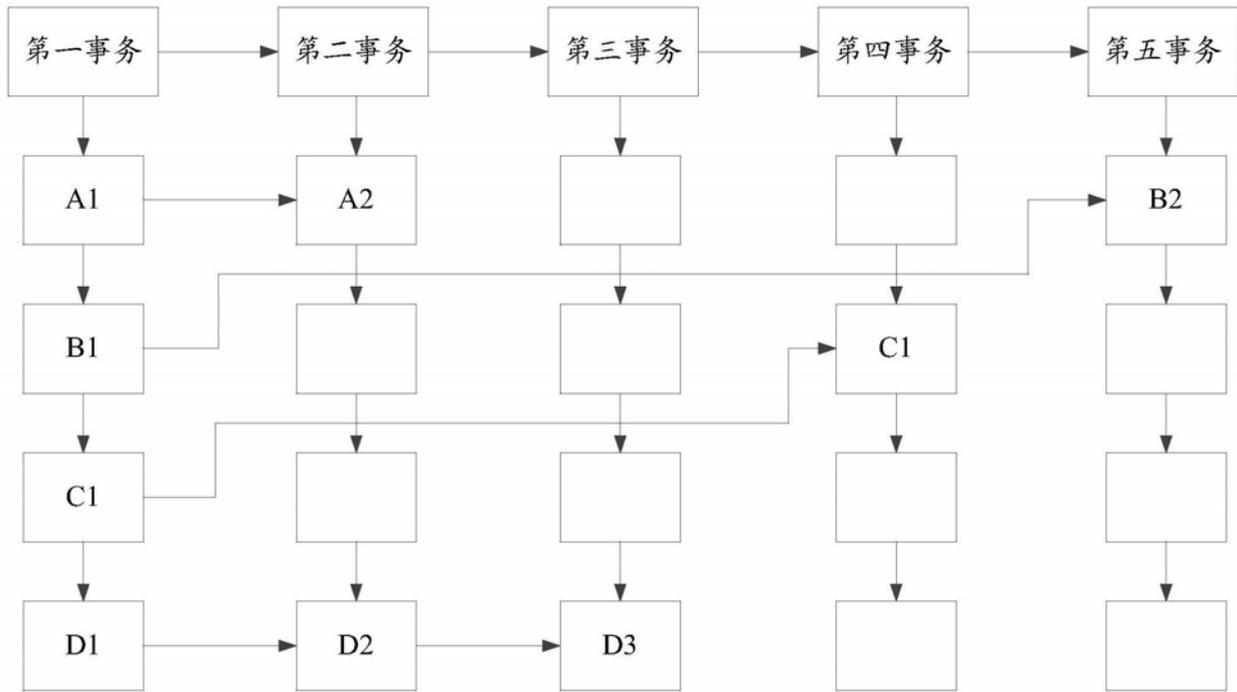


图3

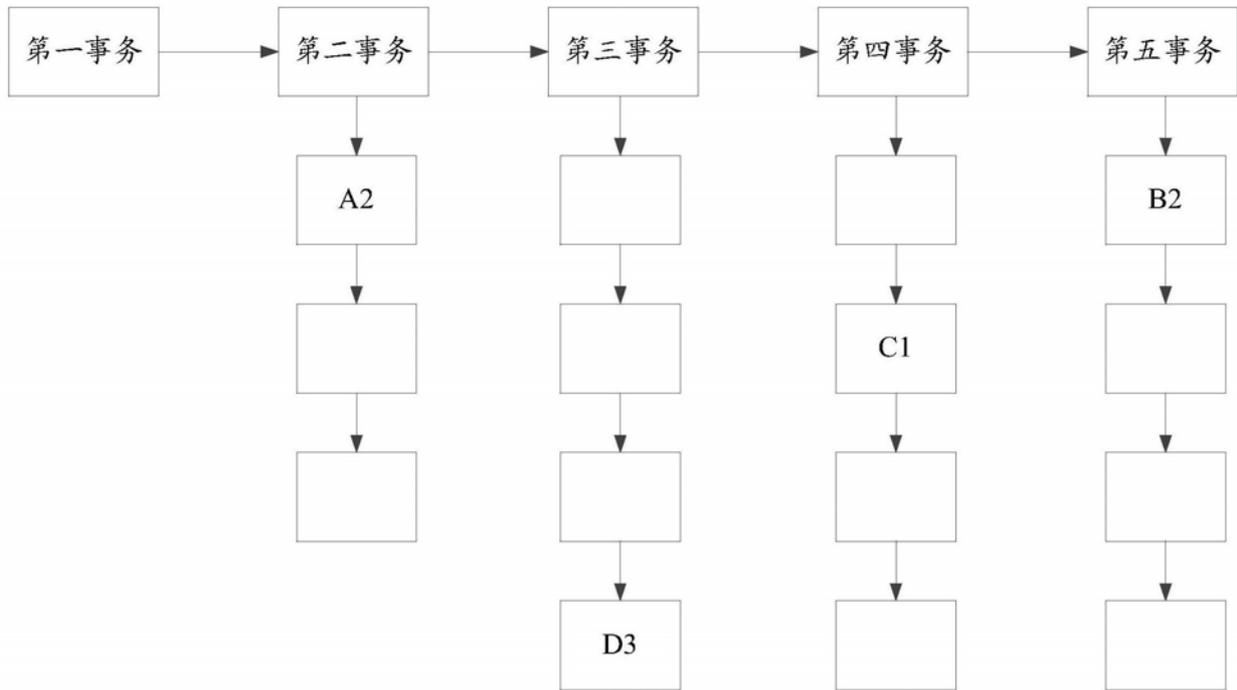


图4

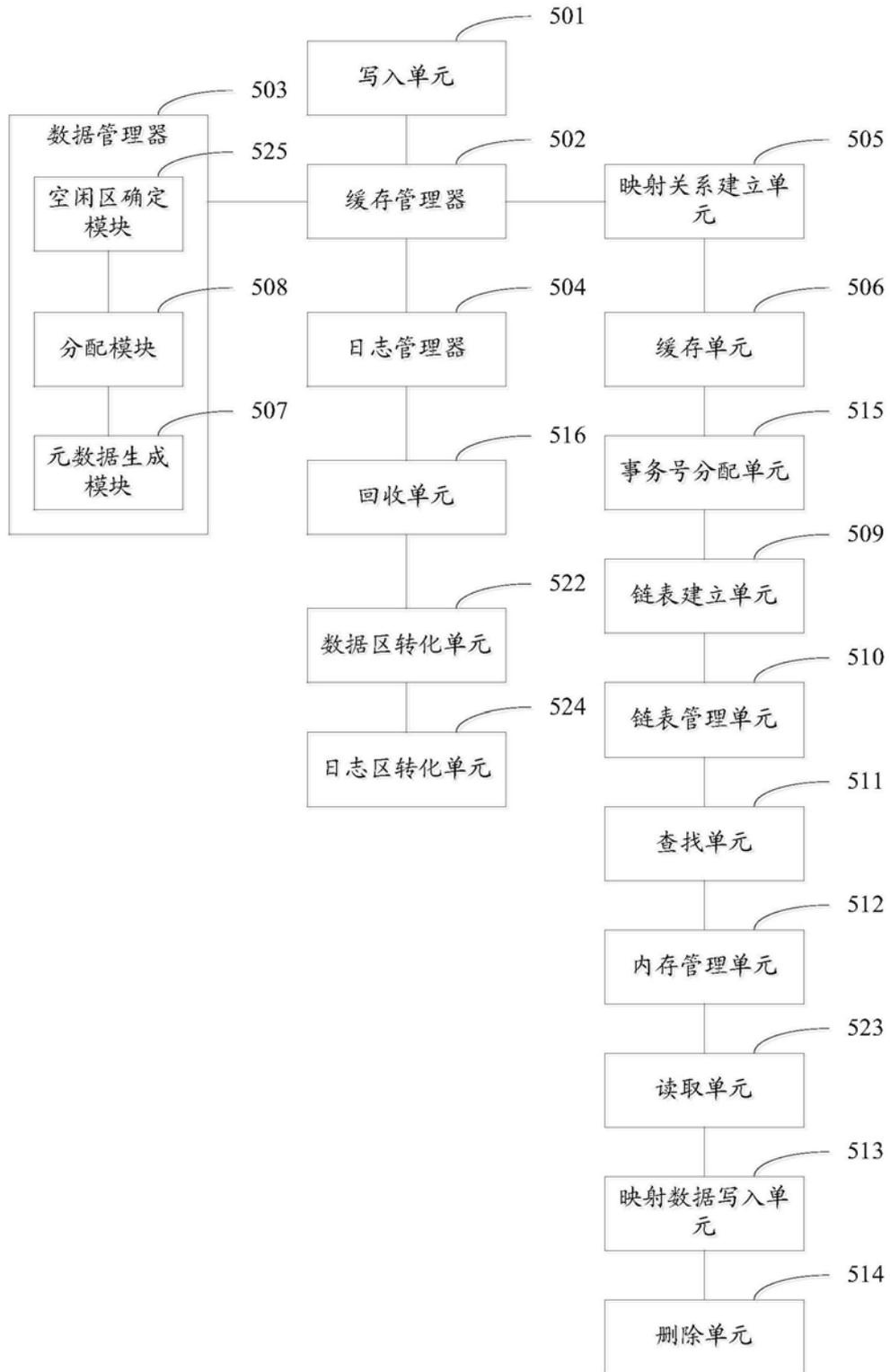


图5

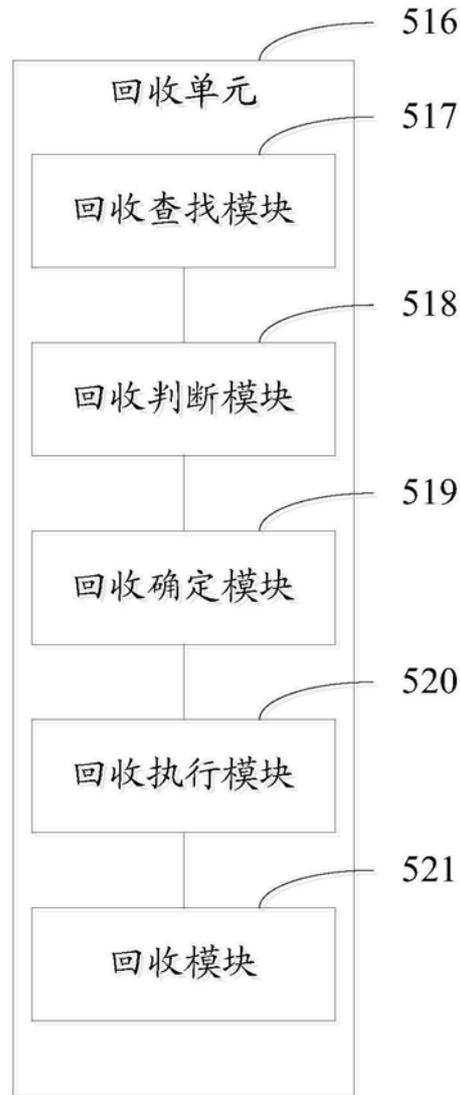


图6

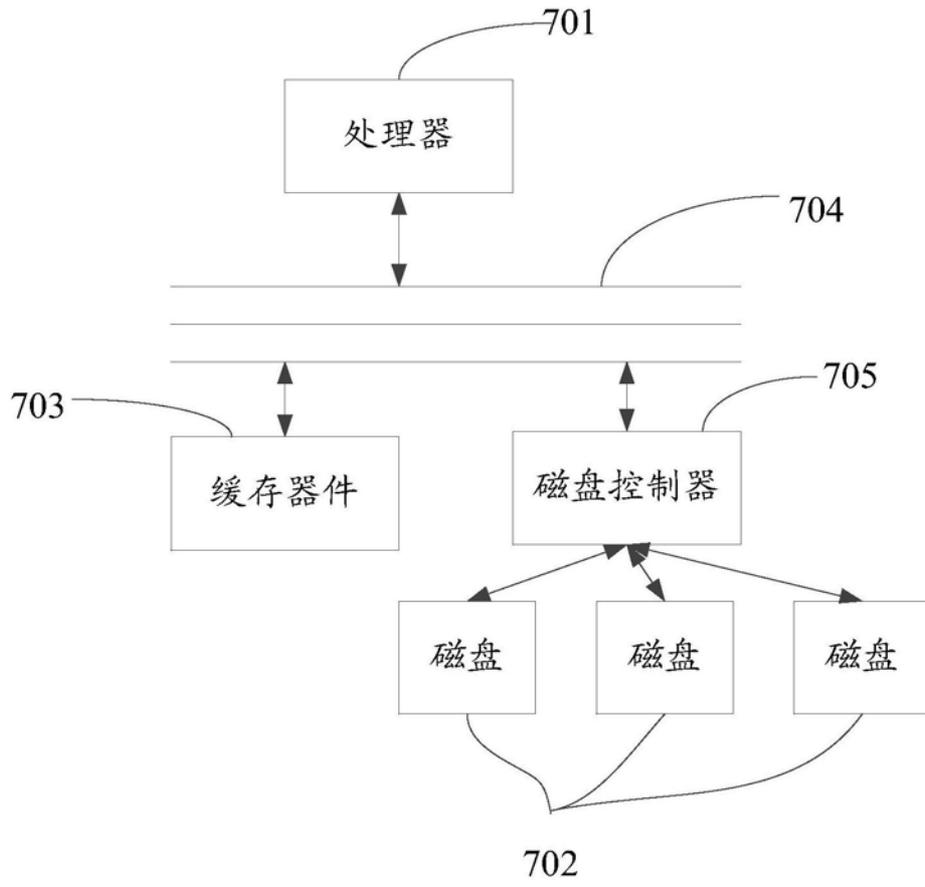


图7