



(12) 发明专利申请

(10) 申请公布号 CN 117008845 A

(43) 申请公布日 2023. 11. 07

(21) 申请号 202311259647.1

(22) 申请日 2023.09.27

(71) 申请人 苏州元脑智能科技有限公司

地址 215000 江苏省苏州市吴中经济开发区郭巷街道官浦路1号9幢

(72) 发明人 刘全利 李志刚 何秋明 云鹏

(74) 专利代理机构 北京康信知识产权代理有限公司 11240

专利代理师 刘晓燕

(51) Int. Cl.

G06F 3/06 (2006.01)

权利要求书4页 说明书16页 附图9页

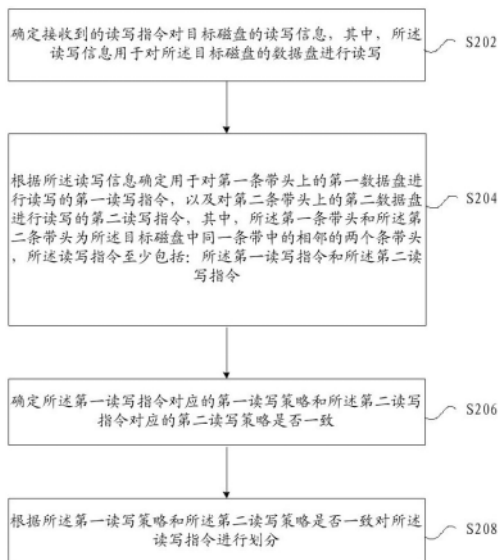
(54) 发明名称

读写指令的划分方法、装置、存储介质及电子装置

(57) 摘要

本申请实施例提供了一种读写指令的划分方法、装置、存储介质及电子装置,其中,该方法包括:确定接收到的读写指令对目标磁盘的读写信息,其中,读写信息用于对目标磁盘的数据盘进行读写;根据读写信息确定用于对第一条带头上的第一数据盘进行读写的第一读写指令,以及对第二条带头上的第二数据盘进行读写的第二读写指令,其中,第一条带头和第二条带头为目标磁盘中同一条带中的相邻的两个条带头,读写指令至少包括:第一读写指令和第二读写指令;确定第一读写指令对应的第一读写策略和第二读写指令对应的第二读写策略是否一致;根据第一读写策略和第二读写策略是否一致对读写指令进行划分。

CN 117008845 A



1. 一种读写指令的划分方法,其特征在于,包括:

确定接收到的读写指令对目标磁盘的读写信息,其中,所述读写信息用于对所述目标磁盘的数据盘进行读写;

根据所述读写信息确定用于对第一条带头上的第一数据盘进行读写的第一读写指令,以及对第二条带头上的第二数据盘进行读写的第二读写指令,其中,所述第一条带头和所述第二条带头为所述目标磁盘中同一条带中的相邻的两个条带头,所述读写指令至少包括:所述第一读写指令和所述第二读写指令;

确定所述第一读写指令对应的第一读写策略和所述第二读写指令对应的第二读写策略是否一致;

根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分。

2. 根据权利要求1所述的方法,其特征在于,根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分,包括:

在所述第一读写策略和所述第二读写策略一致的情况下,将所述第一读写指令和所述第二读写指令划分为同一子读写指令;

在所述第一读写策略和所述第二读写策略不一致的情况下,将所述第一读写指令和所述第二读写指令划分为不同的子读写指令。

3. 根据权利要求1所述的方法,其特征在于,根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分之后,所述方法还包括:

确定所述读写指令对应的子读写指令集合;

确定所述子读写指令集合中的每一子读写指令对应的第三读写策略。

4. 根据权利要求3所述的方法,其特征在于,在所述读写指令为读指令的情况下,确定所述子读写指令集合中的每一子读写指令对应的第三读写策略,包括:

确定所有所述子读写指令对应的第三数据盘是否为故障盘;

根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略。

5. 根据权利要求4所述的方法,其特征在于,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略,包括:

在所有所述第三数据盘均不为故障盘的情况下,确定所述第三读写策略为直接读策略;

在任一所述第三数据盘为故障盘的情况下,确定所述第三读写策略为重构读策略。

6. 根据权利要求4所述的方法,其特征在于,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略之后,所述方法还包括:

确定所述子读写指令集合中的子读写指令的数量;

确定所述子读写指令的数量与预设数量的第一大小关系;

根据所述第一大小关系确定所述第三读写策略对应的执行方式。

7. 根据权利要求6所述的方法,其特征在于,根据所述第一大小关系确定所述第三读写策略对应的执行方式,包括:

在所述第一大小关系指示所述数量等于所述预设数量的情况下,直接执行所述第三读写策略;

在所述第一大小关系指示所述数量大于所述预设数量的情况下,获取所述条带在磁盘中的第一源地址,并将所述第一源地址存储至所述磁盘对应的芯片上的第一存储空间;根据所述第一源地址确定每个所述子读写指令对应的第二源地址,将每个所述第二源地址存储至不同的第二存储空间;根据所述第二存储空间中的第二源地址执行所述第三读写策略。

8. 根据权利要求7所述的方法,其特征在于,在所述第三读写策略为直接读策略的情况下,根据所述第二存储空间中的第二源地址执行所述第三读写策略,包括:

根据所述第二源地址确定所述子读写指令对应的第三条带头;

响应所述子读写指令,以在所述第三条带头中读取所述第三数据盘中的数据。

9. 根据权利要求7所述的方法,其特征在于,在所述第三读写策略为重构读策略的情况下,根据所述第二存储空间中的第二源地址执行所述第三读写策略,包括:

根据所述第二源地址确定所述子读写指令对应的第三条带头;

读取所述第三条带头中第四数据盘中的第一数据和所述第三条带头中校验盘的第二数据,其中,所述第四数据盘为所述第三条带头中除所述第三数据盘之外的数据盘;

根据所述第一数据和所述第二数据确定所述子读写指令对应的数据盘对应的数据。

10. 根据权利要求3所述的方法,其特征在于,在所述读写指令为写指令的情况下,确定所述子读写指令集合中的每一子读写指令对应的第三读写策略,包括:

确定所有所述子读写指令对应的第三数据盘是否为故障盘;

根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略。

11. 根据权利要求10所述的方法,其特征在于,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略,包括:

在所有所述第三数据盘均不为所述故障盘的情况下,根据所述条带中的故障盘的类型确定所述每一子读写指令对应的第三读写策略;

在任一所述第三数据盘为所述故障盘的情况下,确定所述第三读写策略为重构写策略。

12. 根据权利要求11所述的方法,其特征在于,根据所述条带中的故障盘的类型确定所述每一子读写指令对应的第三读写策略,包括:

在所述条带中的故障盘的类型为数据盘的情况下,确定所述第三读写策略为读修改写策略;

在所述条带中的故障盘的类型为校验盘的情况下,确定所述第三读写策略为直接写策略。

13. 根据权利要求10所述的方法,其特征在于,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略之后,所述方法还包括:

确定所述子读写指令集合中的子读写指令的数量;

确定所述子读写指令的数量与预设数量的第二大小关系;

根据所述第二大小关系确定所述第三读写策略的执行方式。

14. 根据权利要求13所述的方法,其特征在于,根据所述第二大小关系确定所述第三读写策略的执行方式,包括:

在所述第二大小关系指示所述数量等于预设数量的情况下,直接执行所述第三读写策略;

在所述第二大小关系指示所述数量大于预设数量的情况下,获取每个所述子读写指令对应的第四条带头中存储的第四数据,并将每个所述第四数据存储至所述磁盘对应的芯片上的不同的第三存储空间;对所述第三存储空间上的第四数据执行所述第三读写策略。

15. 根据权利要求14所述的方法,其特征在于,在所述第三读写策略为重构写策略的情况下,对所述第三存储空间上的第三数据执行所述第三读写策略,包括:

获取所述第三存储空间中的第五数据,其中,所述第五数据为所述第四条带头中第五数据盘中存储的数据,所述第五数据盘为所述第四条带头中除所述第三数据盘之外的数据盘;

获取所述子读写指令中携带的待写入数据;

根据所述第五数据和所述待写入数据计算校验数据;

将所述第三数据盘对应的数据更新为所述待写入数据,以及将所述第四条带头中校验盘对应的数据更新为所述校验数据。

16. 根据权利要求14所述的方法,其特征在于,在所述第三读写策略为读修改写策略的情况下,根据所述第二大小关系确定所述第三读写策略的执行方式,包括:

获取所述第三存储空间中的第五数据和第六数据,其中,所述第五数据为所述第四条带头中的第五数据盘中存储的数据,所述第五数据盘为所述第四条带头中除所述第三数据盘之外的数据盘,所述第六数据为所述第四条带头中的校验盘中存储的数据;

获取所述子读写指令中携带的待写入数据;

根据所述第五数据、所述第六数据和所述待写入数据计算校验数据;

将所述第三数据盘对应的数据更新为所述待写入数据,以及将所述第六数据更新为所述校验数据。

17. 根据权利要求14所述的方法,其特征在于,在所述第三读写策略为直接写策略的情况下,根据所述第二大小关系确定所述第三读写策略的执行方式,包括:

获取所述子读写指令中携带的待写入数据;

将所述第三数据盘对应的数据更新为所述待写入数据,并禁止对所述第四条带头中校验盘对应的数据进行更新。

18. 根据权利要求14所述的方法,其特征在于,对所述第三存储空间上的第三数据执行所述第三读写策略之后,所述方法还包括:

确定所述第三存储空间上的更新后的第三数据;

确定所述第三存储空间对应的条带头;

将所述更新后的第三数据存储至所述第三存储空间对应的条带头中。

19. 根据权利要求1所述的方法,其特征在于,确定接收到的读写指令对目标磁盘的读写信息之后,所述方法还包括:

根据所述读写信息确定所述读写指令对所述目标磁盘的一个或者多个条带进行读写操作;

在确定所述读写指令对多个条带进行读写操作的情况下,将所述读写指令划分为多个目标读写指令,其中,任一所述目标读写指令用于对任一条带进行读写操作,任一所述目标

读写指令包括:所述第一读写指令和所述第二读写指令。

20. 一种读写指令的划分装置,其特征在于,包括:

第一确定模块,用于确定接收到的读写指令对目标磁盘的读写信息,其中,所述读写信息用于对所述目标磁盘的数据盘进行读写;

第二确定模块,用于根据所述读写信息确定用于对第一条带头上的第一数据盘进行读写的第一读写指令,以及对第二条带头上的第二数据盘进行读写的第二读写指令,其中,所述第一带头和所述第二带头为所述目标磁盘中同一条带中的相邻的两个条带头,所述读写指令至少包括:所述第一读写指令和所述第二读写指令;

第三确定模块,用于确定所述第一读写指令对应的第一读写策略和所述第二读写指令对应的第二读写策略是否一致;

划分模块,用于根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分。

21. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质中存储有计算机程序,其中,所述计算机程序被处理器执行时实现所述权利要求1至19任一项中所述的方法的步骤。

22. 一种电子设备,包括存储器、处理器以及存储在所述存储器上并可在所述处理器上运行的计算机程序,其特征在于,所述处理器执行所述计算机程序时实现所述权利要求1至19任一项中所述的方法的步骤。

## 读写指令的划分方法、装置、存储介质及电子装置

### 技术领域

[0001] 本申请实施例涉及通信领域,具体而言,涉及一种读写指令的划分方法、装置、存储介质及电子装置方法及装置。

### 背景技术

[0002] 在存储领域中,通过磁盘阵列自身的冗余性将数据直接或间接存储在多个单独的磁盘上,以达到当一个或多个磁盘发生故障时,数据不丢失的目的,即实现了数据容错。当一个或多个磁盘故障时,为了恢复磁盘阵列的容错,通常采用热备盘替换故障盘进行重构。重构过程中,磁盘阵列被分为多个区域,重构未开始区域、重构进行区域、重构完成区。在热备盘替换故障盘前,整个磁盘阵列处于降级状态。另外对于重建未开始区域,磁盘阵列也是处于降级状态。处于降级状态的磁盘阵列,主机的读写指令IO如果不能被正常处理,不仅导致该IO的用户数据丢失,甚至还会导致磁盘阵列之前存储的数据丢失,造成不可估量的损失。所以可以正常安全地处理分布在降级状态的磁盘阵列上的主机IO,尤为重要。

[0003] 常用的具备冗余功能的磁盘阵列类型包括RAID (Redundant Arrays of Independent Disks,磁盘阵列) 1、5、6、10、50、60,其中RAID1和RAID10,数据被镜像存储于成对的硬盘上,当成对的硬盘中的一个出现故障时,可以对镜像盘进行读和写,不涉及降级的读写。RAID5允许一个盘出现故障时降级为RAID0进行读写。RAID6允许一个盘出现故障时降级为RAID5进行读写,两个盘出现故障时降级为RAID0进行读写。RAID50和RAID60,可以看成由一组子RAID5和RAID6组成,降级读写的方法同RAID5和RAID6。

[0004] 其中,对于RAID5和RAID6降级读写,现有技术中,通常是在RAID条带化技术的基础上,进一步将Stripe (条带)上的IO切分为一个个的Stripe Head (条带头)对应的子IO进行处理,进而对读写指令的拆分次数较多。

[0005] 针对现有技术中,对读写指令的拆分次数较多,导致读写指令的执行效率降低,响应读写指令的延时增大,IOPS降低的问题并未得到有效解决。

### 发明内容

[0006] 本申请实施例提供了一种读写指令的划分方法、装置、存储介质及电子装置,以至少解决相关技术中对读写指令的拆分次数较多,导致读写指令的执行效率降低,响应读写指令的延时增大,IOPS降低的问题。

[0007] 根据本申请的一个实施例,提供了一种读写指令的划分方法,包括:

[0008] 确定接收到的读写指令对目标磁盘的读写信息,其中,所述读写信息用于对所述目标磁盘的数据盘进行读写;根据所述读写信息确定用于对第一条带头上的第一数据盘进行读写的第一读写指令,以及对第二条带头上的第二数据盘进行读写的第二读写指令,其中,所述第一条带头和所述第二条带头为所述目标磁盘中同一条带中的相邻的两个条带头,所述读写指令至少包括:所述第一读写指令和所述第二读写指令;确定所述第一读写指令对应的第一读写策略和所述第二读写指令对应的第二读写策略是否一致;根据所述第一

读写策略和所述第二读写策略是否一致对所述读写指令进行划分。

[0009] 在一个示例性实施例中,根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分,包括:在所述第一读写策略和所述第二读写策略一致的情况下,将所述第一读写指令和所述第二读写指令划分为同一子读写指令;在所述第一读写策略和所述第二读写策略不一致的情况下,将所述第一读写指令和所述第二读写指令划分为不同的子读写指令。

[0010] 在一个示例性实施例中,根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分之后,所述方法还包括:确定所述读写指令对应的子读写指令集合;确定所述子读写指令集合中的每一子读写指令对应的第三读写策略。

[0011] 在一个示例性实施例中,在所述读写指令为读指令的情况下,确定所述子读写指令集合中的每一子读写指令对应的第三读写策略,包括:确定所有所述子读写指令对应的第三数据盘是否为故障盘;根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略。

[0012] 在一个示例性实施例中,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略,包括:在所有所述第三数据盘均不为故障盘的情况下,确定所述第三读写策略为直接读策略;在任一所述第三数据盘为故障盘的情况下,确定所述第三读写策略为重构读策略。

[0013] 在一个示例性实施例中,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略之后,所述方法还包括:确定所述子读写指令集合中的子读写指令的数量;确定所述子读写指令的数量与预设数量的第一大小关系;根据所述第一大小关系确定所述第三读写策略对应的执行方式。

[0014] 在一个示例性实施例中,根据所述第一大小关系确定所述第三读写策略对应的执行方式,包括:在所述第一大小关系指示所述数量等于所述预设数量的情况下,直接执行所述第三读写策略;在所述第一大小关系指示所述数量大于所述预设数量的情况下,获取所述条带在磁盘中的第一源地址,并将所述第一源地址存储至所述磁盘对应的芯片上的第一存储空间;根据所述第一源地址确定每个所述子读写指令对应的第二源地址,将每个所述第二源地址存储至不同的第二存储空间;根据所述第二存储空间中的第二源地址执行所述第三读写策略。

[0015] 在一个示例性实施例中,在所述第三读写策略为直接读策略的情况下,根据所述第二存储空间中的第二源地址执行所述第三读写策略,包括:根据所述第二源地址确定所述子读写指令对应的第三条带头;响应所述子读写指令,以在所述第三条带头中读取所述第三数据盘中的数据。

[0016] 在一个示例性实施例中,在所述第三读写策略为重构读策略的情况下,根据所述第二存储空间中的第二源地址执行所述第三读写策略,包括:根据所述第二源地址确定所述子读写指令对应的第三条带头;读取所述第三条带头中第四数据盘中的第一数据和所述第三条带头中校验盘的第二数据,其中,所述第四数据盘为所述第三条带头中除所述第三数据盘之外的数据盘;根据所述第一数据和所述第二数据确定所述子读写指令对应的数据盘对应的数据。

[0017] 在一个示例性实施例中,在所述读写指令为写指令的情况下,确定所述子读写指

令集合中的每一子读写指令对应的第三读写策略,包括:确定所有所述子读写指令对应的第三数据盘是否为故障盘;根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略。

[0018] 在一个示例性实施例中,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略,包括:在所有所述第三数据盘均不为所述故障盘的情况下,根据所述条带中的故障盘的类型确定所述每一子读写指令对应的第三读写策略;在任一所述第三数据盘为所述故障盘的情况下,确定所述第三读写策略为重构写策略。

[0019] 在一个示例性实施例中,根据所述条带中的故障盘的类型确定所述每一子读写指令对应的第三读写策略,包括:在所述条带中的故障盘的类型为数据盘的情况下,确定所述第三读写策略为读修改写策略;在所述条带中的故障盘的类型为校验盘的情况下,确定所述第三读写策略为直接写策略。

[0020] 在一个示例性实施例中,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略之后,所述方法还包括:确定所述子读写指令集合中的子读写指令的数量;确定所述子读写指令的数量与预设数量的第二大小关系;根据所述第二大小关系确定所述第三读写策略的执行方式。

[0021] 在一个示例性实施例中,根据所述第二大小关系确定所述第三读写策略的执行方式,包括:在所述第二大小关系指示所述数量等于预设数量的情况下,直接执行所述第三读写策略;在所述第二大小关系指示所述数量大于预设数量的情况下,获取每个所述子读写指令对应的第四条带头中存储的第四数据,并将每个所述第四数据存储至所述磁盘对应的芯片上的不同的第三存储空间;对所述第三存储空间上的第四数据执行所述第三读写策略。

[0022] 在一个示例性实施例中,在所述第三读写策略为重构写策略的情况下,对所述第三存储空间上的第三数据执行所述第三读写策略,包括:获取所述第三存储空间中的第五数据,其中,所述第五数据为所述第四条带头中第五数据盘中存储的数据,所述第五数据盘为所述第四条带头中除所述第三数据盘之外的数据盘;获取所述子读写指令中携带的待写入数据;根据所述第五数据和所述待写入数据计算校验数据;将所述第三数据盘对应的数据更新为所述待写入数据,以及将所述第四条带头中校验盘对应的数据更新为所述校验数据。

[0023] 在一个示例性实施例中,在所述第三读写策略为读修改写策略的情况下,根据所述第二大小关系确定所述第三读写策略的执行方式,包括:获取所述第三存储空间中的第五数据和第六数据,其中,所述第五数据为所述第四条带头中的第五数据盘中存储的数据,所述第五数据盘为所述第四条带头中除所述第三数据盘之外的数据盘,所述第六数据为所述第四条带头中的校验盘中存储的数据;获取所述子读写指令中携带的待写入数据;根据所述第五数据、所述第六数据和所述待写入数据计算校验数据;将所述第三数据盘对应的数据更新为所述待写入数据,以及将所述第六数据更新为所述校验数据。

[0024] 在一个示例性实施例中,在所述第三读写策略为直接写策略的情况下,根据所述第二大小关系确定所述第三读写策略的执行方式,包括:获取所述子读写指令中携带的待写入数据;将所述第三数据盘对应的数据更新为所述待写入数据,并禁止对所述第四条带头中校验盘对应的数据进行更新。



[0025] 在一个示例性实施例中,对所述第三存储空间上的第三数据执行所述第三读写策略之后,所述方法还包括:确定所述第三存储空间上的更新后的第三数据;确定所述第三存储空间对应的条带头;将所述更新后的第三数据存储至所述第三存储空间对应的条带头中。

[0026] 在一个示例性实施例中,确定接收到的读写指令对目标磁盘的读写信息之后,所述方法还包括:根据所述读写信息确定所述读写指令对所述目标磁盘的一个或者多个条带进行读写操作;在确定所述读写指令对多个条带进行读写操作的情况下,将所述读写指令划分为多个目标读写指令,其中,任一所述目标读写指令用于对任一条带进行读写操作,任一所述目标读写指令包括:所述第一读写指令和所述第二读写指令。

[0027] 根据本申请的另一个实施例,提供了一种读写指令的划分装置,包括:

[0028] 第一确定模块,用于确定接收到的读写指令对目标磁盘的读写信息,其中,所述读写信息用于对所述目标磁盘的数据盘进行读写;第二确定模块,用于根据所述读写信息确定用于对第一条带头上的第一数据盘进行读写的第一读写指令,以及对第二条带头上的第二数据盘进行读写的第二读写指令,其中,所述第一条带头和所述第二条带头为所述目标磁盘中同一条带中的相邻的两个条带头,所述读写指令至少包括:所述第一读写指令和所述第二读写指令;第三确定模块,用于确定所述第一读写指令对应的第一读写策略和所述第二读写指令对应的第二读写策略是否一致;划分模块,用于根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分。

[0029] 根据本申请的又一个实施例,还提供了一种计算机可读存储介质,所述计算机可读存储介质中存储有计算机程序,其中,所述计算机程序被设置为运行时执行上述任一项方法实施例中的步骤。

[0030] 根据本申请的又一个实施例,还提供了一种电子设备,包括存储器和处理器,所述存储器中存储有计算机程序,所述处理器被设置为运行所述计算机程序以执行上述任一项方法实施例中的步骤。

[0031] 通过本申请,由于确定接收到的用于对目标磁盘的数据盘进行读写的读写信息;根据读写信息确定第一读写指令以及第二读写指令,其中,第一读写指令用于对第一条带头上的第一数据盘进行读写,第二读写指令用于对第二条带头上的第二数据盘进行读写,第一条带头和第二条带头为目标磁盘中同一条带中的相邻的两个条带头;根据判断第一读写指令对应的第一读写策略与第二读写指令对应的第二读写策略是否一致划分读写指令。因此,可以解决现有技术中,对读写指令的拆分次数较多,导致读写指令的执行效率降低,响应读写指令的延时增大,IOPS降低的问题,进而达到提升读写指令的执行效率的效果。

## 附图说明

[0032] 图1是本申请实施例的一种读写指令的划分方法的计算机终端的硬件结构框图;

[0033] 图2是根据本申请实施例的一种读写指令的划分的流程图,如图1所示;

[0034] 图3是现有技术中的读写指令的划分方案示意图;

[0035] 图4是根据本申请实施例的Stripe内IO不拆分的读写策略;

[0036] 图5是根据本申请实施例的IO拆分的整体流程图;

[0037] 图6是根据本申请实施例的IO按Stripe拆分示意图(一);

- [0038] 图7是根据本申请实施例的IO按Stripe拆分示意图(二);
- [0039] 图8是根据本申请实施例的IO在Stripe内的不拆分示意图;
- [0040] 图9是根据本申请实施例的Stripe被拆分的读IO的执行流程图;
- [0041] 图10是根据本申请实施例的Stripe被拆分的写IO的执行流程图;
- [0042] 图11是根据本申请实施例中的RAID5故障读示意图;
- [0043] 图12是根据本申请实施例的RAID5故障写示意图;
- [0044] 图13是根据本申请实施例的读写指令的划分装置的结构框图。

### 具体实施方式

[0045] 下文中将参考附图并结合实施例来详细说明本申请的实施例。

[0046] 需要说明的是,本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。

[0047] 本申请实施例中所提供的方法实施例可以在移动终端、计算机终端或者类似的运算装置中执行。以运行在计算机终端上为例,图1是本申请实施例的一种读写指令的划分方法的计算机终端的硬件结构框图。如图1所示,计算机终端可以包括一个或多个(图1中仅示出一个)处理器102(处理器102可以包括但不限于微处理器MCU或可编程逻辑器件FPGA等的处理装置)和用于存储数据的存储器104,其中,上述计算机终端还可以包括用于通信功能的传输设备106以及输入输出设备108。本领域普通技术人员可以理解,图1所示的结构仅为示意,其并不对上述计算机终端的结构造成限定。例如,计算机终端还可包括比图1中所示更多或者更少的组件,或者具有与图1所示不同的配置。

[0048] 存储器104可用于存储计算机程序,例如,应用软件的软件程序以及模块,如本申请实施例中的读写指令的划分方法对应的计算机程序,处理器102通过运行存储在存储器104内的计算机程序,从而执行各种功能应用以及数据处理,即实现上述的方法。存储器104可包括高速随机存储器,还可包括非易失性存储器,如一个或者多个磁性存储装置、闪存、或者其他非易失性固态存储器。在一些实例中,存储器104可进一步包括相对于处理器102远程设置的存储器,这些远程存储器可以通过网络连接至计算机终端。上述网络的实例包括但不限于互联网、企业内部网、局域网、移动通信网及其组合。

[0049] 传输设备106用于经由一个网络接收或者发送数据。上述的网络具体实例可包括计算机终端的通信供应商提供的无线网络。在一个实例中,传输设备106包括一个网络适配器(Network Interface Controller,简称为NIC),其可通过基站与其他网络设备相连从而可与互联网进行通讯。在一个实例中,传输设备106可以为射频(Radio Frequency,简称为RF)模块,其用于通过无线方式与互联网进行通讯。

[0050] 图2是根据本申请实施例的一种读写指令的划分的流程图,如图2所示,该流程包括如下步骤:

[0051] 步骤S202,确定接收到的读写指令对目标磁盘的读写信息,其中,所述读写信息用于对所述目标磁盘的数据盘进行读写;

[0052] 步骤S204,根据所述读写信息确定用于对第一条带头上的第一数据盘进行读写的第一读写指令,以及对第二条带头上的第二数据盘进行读写的第二读写指令,其中,所述第一条带头和所述第二条带头为所述目标磁盘中同一条带中的相邻的两个条带头,所述读写

指令至少包括:所述第一读写指令和所述第二读写指令;

[0053] 步骤S206,确定所述第一读写指令对应的第一读写策略和所述第二读写指令对应的第二读写策略是否一致;

[0054] 可以理解的是,上述读写策略可以包括但不限于:

[0055] RCW(Reconstruct Write,重构写),通常被称为大写。通过读出本次写入磁盘阵列的读写指令的位置以外的其他位置的数据,和本次读写指令的新数据,计算出新的校验数据的方法;

[0056] RMW(Read Modify Write,读修改写),通常被称为小写。通过读出本次写入磁盘阵列的读写指令的位置的老数据、老的校验数据、和读写指令的新数据,计算出新的校验数据的方法;

[0057] FSW(Full Stripe Write,满条带写),通常被称为满写。写入磁盘阵列的数据量正好等于条带大小,如此一来便不需要额外的读旧数据的操作就能计算出新的校验数据的方法;

[0058] RCR(Reconstruct Read,重构读)。通过读出本次读写指令的位置以外的其他位置的数据、校验盘的数据,计算恢复出本次读写指令涉及的数据方法。

[0059] 步骤S208,根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分。

[0060] 通过上述步骤,由于确定接收到的用于对目标磁盘的数据盘进行读写的读写信息;根据读写信息确定第一读写指令以及第二读写指令,其中,第一读写指令用于对第一条带头上的第一数据盘进行读写,第二读写指令用于对第二条带头上的第二数据盘进行读写,第一条带头和第二条带头为目标磁盘中同一条带中的相邻的两个条带头;根据判断第一读写指令对应的第一读写策略与第二读写指令对应的第二读写策略是否一致划分读写指令。因此,可以解决现有技术中,对读写指令的拆分次数较多,导致读写指令的执行效率降低,响应读写指令的延时增大,IOPS降低的问题,进而达到提升读写指令的执行效率的效果。

[0061] 可选的,上述步骤S208包括以下两种情况:在所述第一读写策略和所述第二读写策略一致的情况下,将所述第一读写指令和所述第二读写指令划分为同一子读写指令;在所述第一读写策略和所述第二读写策略不一致的情况下,将所述第一读写指令和所述第二读写指令划分为不同的子读写指令。

[0062] 可以理解的是,可以根据运用的读写策略是否相同确定对读写指令的划分,即运用的读写策略相同,则将相同的读写策略对应的读写指令划分为同一子读写指令;而运用的读写策略不同,则将不同的读写策略对应的读写指令划分为不同的子读写指令。

[0063] 举例说明,在第一读写策略和第二读写策略都是RCW的策略的情况下,那么可以将第一读写策略对应的第一读写指令以及第二读写策略对应的第二读写指令划分为同一个子读写指令;在第一读写策略为RCR的策略,而第二读写策略为RMW的策略的情况下,那么可以将第一读写策略对应的第一读写指令以及第二读写策略对应的第二读写指令划分为不同的子读写指令。

[0064] 根据上述实施例,根据分布于相邻的两个条带头上的读写指令对应的读写策略是否一致对读写指令进行划分,有效地降低了读写指令的拆分次数,进而提升了读写指令的

执行效率。

[0065] 可选的,根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分之后,还需要执行以下步骤:确定所述读写指令对应的子读写指令集合;确定所述子读写指令集合中的每一子读写指令对应的第三读写策略。

[0066] 由于读写指令分为:读指令和写指令,因此,在所述读写指令为读指令的情况下,确定所述子读写指令集合中的每一子读写指令对应的第三读写策略的方式如下:

[0067] 确定所有所述子读写指令对应的第三数据盘是否为故障盘;根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略。

[0068] 可以理解的是,在所述读写指令为读指令的情况下,读写策略的确定受数据盘是否为故障盘的影响,可以根据数据盘是否是故障盘确定每一子读写指令对应的读写策略。

[0069] 具体的,在所有所述第三数据盘均不为故障盘的情况下,确定所述第三读写策略为直接读策略;在任一所述第三数据盘为故障盘的情况下,确定所述第三读写策略为重构读策略。

[0070] 可以理解的是,在读写指令为读指令,且数据盘是故障盘的情况下,需要采取RCR策略(也就是上述重构读策略)将故障盘中损坏的数据恢复出来,然后读到主机;在数据盘不是故障盘的情况下,则可以采取直接读的方式。

[0071] 可选的,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略之后,所述方法还包括:确定所述子读写指令集合中的子读写指令的数量;确定所述子读写指令的数量与预设数量的第一大小关系;根据所述第一大小关系确定所述第三读写策略对应的执行方式。

[0072] 需要说明的是,上述预设数量可以理解为1,在上述子读写指令的数量为1的情况下,可以理解为没有对读写指令进行拆分;在上述子读写指令的数量大于1的情况下,可以理解为对读写指令进行了拆分。

[0073] 可以理解的是,在根据数据盘是否为故障盘进而确定子读写指令的读写策略之后,要根据子读写指令集合中子读写指令的数量是否为1确定读写策略对应的不同的执行方式。

[0074] 可选的,在所述第一大小关系指示所述数量等于所述预设数量的情况下,即在没有对读写指令进行拆分的情况下,直接执行所述第三读写策略;在所述第一大小关系指示所述数量大于所述预设数量的情况下,即在对读写指令进行了拆分的情况下,获取所述条带在磁盘中的第一源地址,并将所述第一源地址存储至所述磁盘对应的芯片上的第一存储空间;根据所述第一源地址确定每个所述子读写指令对应的第二源地址,将每个所述第二源地址存储至不同的第二存储空间;根据所述第二存储空间中的第二源地址执行所述第三读写策略。

[0075] 在子读写指令集合中仅包含一个子读写指令的情况下,可以直接执行上述子读写指令对应的读写策略。

[0076] 在子读写指令集合中包含N(N为大于1的整数)个子读写指令的情况下,则要获取条带的第一源地址以及根据第一源地址确定子读写指令集合中的每一个子读写指令对应的第二源地址,然后将第一源地址存储至芯片侧的第一存储空间上,例如PRP-List,将不同的第二源地址分别存储到芯片侧的不同的第二存储空间,例如,PRP-SubList,最后根据第

二源地址执行读写策略,其中,第一源地址和第二源地址要对应存储到对应的存储空间中,其中,每一个第二源地址要存储到不同的存储空间中。

[0077] 需要说明的是,第二源地址可以理解为每个子读写指令对应的条带头的源地址,其中,每个子读写指令用于对对应的条带头上的任一数据盘进行数据读取。

[0078] 上述方法的目的是将磁盘的条带的源地址发送到芯片侧,进而重新计算子I0对应的条带头的源地址。

[0079] 需要说明的是,子I0对应的条带头的数量可以是一个也可以是多个。

[0080] 可选的,在所述第三读写策略为直接读策略的情况下,根据所述第二存储空间中的第二源地址执行所述第三读写策略,包括:根据所述第二源地址确定所述子读写指令对应的第三条带头;响应所述子读写指令,以在所述第三条带头中读取所述第三数据盘中的数据。

[0081] 可以理解的是,在子读写指令集合中包含N(N为大于1的整数)个子读写指令,且读写策略为直接读的策略的情况下,要根据第二源地址确定子读写指令对应的第三条带头,进而可以在第三条带头中读取数据盘中的数据。

[0082] 可选的,在所述第三读写策略为重构读策略的情况下,根据所述第二存储空间中的第二源地址执行所述第三读写策略,包括:根据所述第二源地址确定所述子读写指令对应的第三条带头;读取所述第三条带头中第四数据盘中的第一数据和所述第三条带头中校验盘的第二数据,其中,所述第四数据盘为所述第三条带头中除所述第三数据盘之外的数据盘;根据所述第一数据和所述第二数据确定所述子读写指令对应的数据盘对应的数据。

[0083] 也就是说,在第三读写策略为重构读策略的情况下,需要针对该数据所在的stripe Head(条带头)内的数据和校验进行若干次RCR(重构读)计算,恢复出丢失的数据后,再向主机传输。

[0084] 由于读写指令分为:读指令和写指令,因此,在所述读写指令为写指令的情况下,确定所述子读写指令集合中的每一子读写指令对应的第三读写策略,包括:确定所有所述子读写指令对应的第三数据盘是否为故障盘;根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略。

[0085] 可以理解的是,在读写指令为写指令的情况下,同样需要确定子读写指令对应的第三数据盘是否是故障盘,进而根据第三数据盘是否为故障盘确定子读写指令对应的执行方式。

[0086] 具体的,在所有所述第三数据盘均不为所述故障盘的情况下,根据所述条带中的故障盘的类型确定所述每一子读写指令对应的第三读写策略;在任一所述第三数据盘为所述故障盘的情况下,确定所述第三读写策略为重构写策略。

[0087] 可以理解的是,在所有第三数据盘都不是故障盘的情况下,要根据条带中的故障盘的类型确定读写策略;在第三数据盘中有任一故障盘的情况下,则可以采用重构写的策略。

[0088] 可选的,根据所述条带中的故障盘的类型确定所述每一子读写指令对应的第三读写策略,包括:在所述条带中的故障盘的类型为数据盘的情况下,确定所述第三读写策略为读修改写策略;在所述条带中的故障盘的类型为校验盘的情况下,确定所述第三读写策略为直接写策略。

[0089] 可以理解的是,条带中故障盘的类型可以分为数据盘或者故障盘为校验盘两种。其中,在故障盘为数据盘的情况下,采用读修改写的策略,而在故障盘为校验盘的情况下,采用直接写的策略。

[0090] 其中,采用读修改写的策略,要对数据盘和故障盘进行更新;在采用直接写的策略计算完后,只需要对数据盘进行更新,而不需要更新校验盘。

[0091] 可选的,根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略之后,还需要确定所述子读写指令集合中的子读写指令的数量;确定所述子读写指令的数量与预设数量的第二大小关系;根据所述第二大小关系确定所述第三读写策略的执行方式。

[0092] 需要说明的是,上述预设数量可以为1。

[0093] 可以理解的是,根据所有第三数据盘是否为故障盘确定每一子读写指令对应的第三读写策略之后,还需要确定子读写指令集合中的子读写指令的数量是否为1,根据子读写指令的数量是否为1确定相应的执行方式。

[0094] 可选的,根据所述第二大小关系确定所述第三读写策略的执行方式,包括:在所述第二大小关系指示所述数量等于预设数量的情况下,直接执行所述第三读写策略;在所述第二大小关系指示所述数量大于预设数量的情况下,获取每个所述子读写指令对应的第四条带头中存储的第四数据,并将每个所述第四数据存储至所述磁盘对应的芯片上的不同的第三存储空间;对所述第三存储空间上的第四数据执行所述第三读写策略。

[0095] 可以理解的是,在子读写指令集合中的子读写指令的数量为1的情况下,可以直接执行第三读写策略;而在子读写指令集合中的子读写指令的数量为M(M为大于1的整数)的情况下,将每个所述子读写指令对应的第四条带头中存储的第四数据存储至所述磁盘对应的芯片上的不同的第三存储空间,进而对第三存储空间中的数据执行所述第三读写策略。

[0096] 可选的,在所述第三读写策略为重构写策略的情况下,对所述第三存储空间上的第三数据执行所述第三读写策略,包括:获取所述第三存储空间中的第五数据,其中,所述第五数据为所述第四条带头中第五数据盘中存储的数据,所述第五数据盘为所述第四条带头中除所述第三数据盘之外的数据盘;获取所述子读写指令中携带的待写入数据;根据所述第五数据和所述待写入数据计算校验数据;将所述第三数据盘对应的数据更新为所述待写入数据,以及将所述第四条带头中校验盘对应的数据更新为所述校验数据。

[0097] 可以理解的是,在第三读写策略为重构写策略的情况下,要获取上述第三存储空间对应的数据,同时,获取子读写指令中的待写入数据,根据上述第三存储空间对应的数据以及待写入数据计算校验数据,进而将数据盘中的数据更新为待写入数据,将校验盘的数据更新为校验数据。

[0098] 可选的,在所述第三读写策略为读修改写策略的情况下,根据所述第二大小关系确定所述第三读写策略的执行方式,包括:获取所述第三存储空间中的第五数据和第六数据,其中,所述第五数据为所述第四条带头中的第五数据盘中存储的数据,所述第五数据盘为所述第四条带头中除所述第三数据盘之外的数据盘,所述第六数据为所述第四条带头中的校验盘中存储的数据;获取所述子读写指令中携带的待写入数据;根据所述第五数据、所述第六数据和所述待写入数据计算校验数据;将所述第三数据盘对应的数据更新为所述待写入数据,以及将所述第六数据更新为所述校验数据。

[0099] 在第三读写策略为读修改写的策略的情况下,获取第三存储空间中的第五数据盘和校验盘中的数据,分别为第五数据和第六数据,同时,获取子读写指令中的待写入数据,然后根据第五数据和第六数据计算校验数据,同样的,将第三数据盘对应的数据更新为待写入数据,将校验盘对应的数据更新为校验数据。

[0100] 可选的,在所述第三读写策略为直接写策略的情况下,根据所述第二大关系确定所述第三读写策略的执行方式,包括:获取所述子读写指令中携带的待写入数据;将所述第三数据盘对应的数据更新为所述待写入数据,并禁止对所述第四条带头中校验盘对应的数据进行更新。

[0101] 可以理解的是,在第三读写策略为直接写的略的情况下,不需要对校验盘进行更新,只需要对数据盘进行更新,即将第三数据盘对应的数据更新为待写入数据即可。

[0102] 可选的,对所述第三存储空间上的第三数据执行所述第三读写策略之后,确定所述第三存储空间上的更新后的第三数据;确定所述第三存储空间对应的条带头;将所述更新后的第三数据存储至所述第三存储空间对应的条带头中。

[0103] 可以理解的是,可以确定第三存储空间对应的更新后的第三数据以及对应的条带头,将第三数据存储到第三存储空间对应的条带头中。也就是说,在对芯片侧的第三存储空间中的数据进行更新后,根据更新后的数据对第三存储空间对应的磁盘中的条带头中的数据进行更新。

[0104] 可选的,确定接收到的读写指令对目标磁盘的读写信息之后,所述方法还包括:根据所述读写信息确定所述读写指令对所述目标磁盘的一个或者多个条带进行读写操作;在确定所述读写指令对多个条带进行读写操作的情况下,将所述读写指令划分为多个目标读写指令,其中,任一所述目标读写指令用于对任一条带进行读写操作,任一所述目标读写指令包括:所述第一读写指令和所述第二读写指令。

[0105] 可以理解的是,要根据读写信息确定读写操作对应的目标磁盘的一个或多个条带,在确定对多个条带进行读写操作的情况下,要将读写指令划分为不同的目标读写指令。

[0106] 通过以上的实施方式的描述,本领域的技术人员可以清楚地了解到根据上述实施例的方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本申请的技术方案本质上或者说对现有技术做出贡献的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(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,或者网络设备)执行本申请各个实施例所述的方法。

[0107] 为了更好的理解上述读写指令的划分方法的过程,以下再结合可选实施例对上述读写指令的划分方法的实现流程进行说明,但不用于限定本申请实施例的技术方案。

[0108] 在理解读写指令的划分方法的过程之前,需要了解现有技术中的读写指令的划分方案,如图3所示,图3是现有技术中的读写指令的划分方案示意图。

[0109] 对于RAID5和RAID6降级读写,现有技术中,通常是在RAID条带化技术的基础上,进一步将Stripe(条带)上的IO切分为一个个的Stripe Head(条带头)分开进行处理。如下图所示,主机的IO分布为0/1/2/3/4/5,当RAID组中的D0盘故障后,主机IO以Stripe Head为单位,切分为0/4、1/5、2、3的子IO分别进行处理。对于读IO,子IO 0/4,需要先恢复出IO 0,然后读0和4;子IO3可以直接读;子IO 1/5,需要先恢复1,然后读1和5;子IO 2需要先恢复,然

后读。对于写IO,子IO 0/4、子IO 1/5、子IO 2 均可以分别采用大写的方式写盘,子IO 3只能小写的方式写盘。

[0110] 对读写策略相同的连续Stripe Head不切分,从另外一种角度可以有效地降低了磁盘整列降级后的IO拆分次数和读写盘的合并次数,大大提升了磁盘整列地读写效率。

[0111] 本实施例的方案为:

[0112] (1) IO拆分分析计算。首先,将分布在多个Stripe的IO拆分为单个Stripe上的IO进行处理。然后在Stripe内根据IO在降级后的Raid上的分布情况,分析连续的Stripe Head上的IO是否可以采用同一种读写策略,对于可采用同一种读写策略的Stripe Head,直接跳过,不进行拆分。相邻的Stripe Head的读写策略与前一个Stripe Head不同,从相邻的Stripe Head开始拆分。

[0113] (2) 拆分后条带内IO的策略。对于Stripe内不进行拆分的IO,由于读写盘的策略不只是受到因IO大小不同导致读写盘次数不同(比如写小IO采用RMW的方式,写大IO采用RCW,写满条带IO采用FSW)的影响,同时受IO涉及到故障盘的影响,所以会被强制指定为某一种读写策略。图4是根据本申请实施例的Stripe内IO不拆分的读写策略,如图4所示,写IO,如果Raid组中不存在故障的情况下,为了减少读写盘的次数,采用RMW的策略,通过本次新数据0、读出的老数据0、读出的老校验P1,计算出新校验的P1。但是由于Raid组中存在故障盘,只能被强制为RCW的策略,即读出Stripe Head上除校验P1和数据0以外的其他数据,计算新的校验P1。

[0114] 而对于Stripe内拆分后的IO,读写策略同Stripe内部不进行IO拆分的场景,不同的是,读写盘前的准备。写场景,使用DMA先将原始的IO对应的Data从主机分段搬移到每子IO分别对应的存储空间,子IO依据各自的存储空间进行写;读场景,使用DMA,将原始的PRP搬移到PRP存储空间中,软件或者硬件根据拆分的信息计算子IO对应的PRP存储空间,子IO依据各自的存储空间中的PRP分别进行读。

[0115] 以下为本示例性实施例的具体流程:

[0116] (1) IO拆分分析计算:

[0117] 图5是根据本申请实施例的IO拆分的整体流程图,将IO按条带拆分后,对条带内的IO进一步分析,如果连续的Stripe Head上的读写策略不一致,将IO拆分为分布在多个Stripe Head组上的子IO,如图5所示:

[0118] 步骤S501,IO按Stripe拆分;

[0119] 步骤S502,确定IO在连续的Stripe Head上的读写策略是否一致;在一致的情况下,执行步骤S502;在不一致的情况下,执行步骤S503;

[0120] 步骤S503,将IO拆分为分布在多个子条带上的子IO;

[0121] 需要说明的是,子条带通过相邻的条带头构成。

[0122] 步骤S504,判断当前Stripe上的IO拆分是否完成,在完成的情况下,结束;在没有完成的情况下,执行步骤S502。

[0123] 以RAID5为例,图6是根据本申请实施例的IO按Stripe拆分示意图(一),如图6所示,图6中的IO分布在两个Stripe上,被拆分为IO 0/1/2/3/4/5和IO 6/7/8。

[0124] 图7是根据本申请实施例的IO在Stripe内的拆分示意图(二),如图7所示,分布在Stripe内的写IO,Stripe Head2的子IO 0/4可采用“RCW”的策略,Stripe Head3的子IO 1/5



采用“RCW”的策略，Stripe Head0的子IO 2可采用RMW的策略，Stripe Head1的子IO3可采用RMW的策略。以连续的Stripe Head组为单位，将分布在Stripe内的该写IO拆分为子IO 0/1/4/5和子IO 2/3。

[0125] 图8是根据本申请实施例的IO在Stripe内的不拆分示意图，如图8所示，图8中分布在Stripe内的读IO，Stripe Head0-2上的子IO 6、7、8均需要直接读的策略，读的策略相同，所以不对该条带内的IO进行拆分。

[0126] (2) 拆分后条带内IO的策略：

[0127] 1) Stripe内没有被拆分的IO；

[0128] 读IO如果IO没有涉及到故障盘，可以采用直接读的方式。如果IO涉及到故障盘，需要采用“RCR”，将损坏的数据恢复出来，然后读到主机。

[0129] 写IO，IO之外的数据盘出现故障，采用“RMW”的方式计算完后，更新数据盘和校验盘。

[0130] 写IO，IO之外的校验盘出现故障，采用直接写的方式，对校验盘不更新。

[0131] 写IO，IO涉及到故障盘，采用 RCW或者先恢复故障盘数据然后根据IO的大小决定采用RMW、RCW、FSW的方式计算，最后将不涉及故障盘的IO更新到数据盘，将校验数据更新到校验盘，跳过涉及到故障盘的IO。

[0132] 2) Stripe内存在被拆分的IO的执行；

[0133] 由于IO的写盘的策略同Stripe内没有被拆分的IO，这里不详细描述，只针对Stripe内拆分的子IO执行前需要准备的基础进行说明。

[0134] Stripe被拆分的写读IO的执行流程如图9所示，图9是根据本申请实施例的Stripe被拆分的写读IO的执行流程图，其中，A-SubList是为了存储子IO对应的存储空间。

[0135] 步骤S901，DMA部分数据搬移，Stripe内拆分的每个子IO对应的主机数据存储多个A-Sublist中；

[0136] 步骤S902，每个子IO以对应A-Sublist的输入，进行写盘；

[0137] 步骤S903，判断拆分的子IO是否执行完成，在完成的情况下，结束，在没有完成的情况下，转到步骤S902。

[0138] Stripe被拆分的读写IO的执行流程如图10所示。图10是根据本申请实施例的Stripe被拆分的写读IO的执行流程图，10中PRP-List为存储主机PRP的存储空间，PRP-SubList为存储每个子IO对应的主机的PRP的存储空间。

[0139] 步骤S1001，DMA搬移PRP，存储在PRP-List；

[0140] 步骤S1002，根据Strpe内拆分的IO，重定向生成每个子IO对应的PRP-SubList；

[0141] 步骤S1003，每个子IO以对应的PRP-Sublist的输入，进行读盘；

[0142] 步骤S1004，确定拆分的子IO是否完成，在完成的情况下，结束，在没有完成的情况下，执行步骤S1003。

[0143] 为了更清晰地描述本申请的整体流程和方法，下面我们以RAID5 4D+1P组成的RAID5出现一块盘故障的场景，进行逐一说明。

[0144] 1) 图11是根据本申请实施例中的1RAID5一块盘故障读示意图，如图11所示：

[0145] 针对图11中的1、2，该IO没有涉及到故障盘位置，其流程同磁盘阵列正常时的读IO流程。

[0146] 针对图11中的3, I0涉及到了故障盘, I0所在的Stripe Head均可以采用“RCR”的策略, 不需要条带拆分。需要针对该数据所在的Stripe Head内的数据和校验进行若干次“RCR”计算, 恢复出丢失的数据后, 再向主机传输。

[0147] 针对图11中的4、5、6、7, 经过拆条带分析, 拆分为分布在涉及到故障盘的子条带1和未涉及到故障盘的子条带2上的I0\_0和I0\_1(相当于上述实施例中的子读写指令), DMA将原始的PRP搬移到芯片侧PRP-List, 硬件或者软件重新定向后生成子条带的PRP-subList1和PRP-subList2。子I0\_1使用PRP-subList1做“RCR”运算后恢复丢失的数据, 再向主机传输。子I0\_2使用PRP-subList2, 同磁盘阵列正常时的读I0流程。

[0148] 2) 图12是根据本申请实施例的RAID5一块盘故障写示意图, 如图12所示:

[0149] 针对图12中的1、4、7, I0涉及到盘之外的其他数据盘故障, 被强制为“RMW”写策略的写I0处理流程。

[0150] 针对图12中的2、5、8, 校验盘故障的场景, 直接将要写的数据落盘, 不做校验数据更新。

[0151] 图12中的3、9经过条带拆分分析, 每个Stripe Head上的I0均可以采用“RCW”的写策略, 不需要条带拆分。

[0152] 图12中的6经过条带拆分分析, 2个Stripe Head上的4、1/5, 均可采用“RCW”的写策略, 拆分为分布在子条带1的子I0 1/4/5。剩余两个的Stripe Head上的2、3, 均可采用“RMW”的写策略, 拆分为分布在子条带2的子I0 2/3。DMA将主机数据先分段搬移到存储在本地A-subList1(1/4/5)和A-subList2(2/3)中。子I0 1/4/5使用A-subList1完成后续的落盘, 子I0 2/3使用A-subList2完成后续的落盘。

[0153] 通过上述实施例, 可以有以下有益效果:

[0154] (1) 读写拆分次数较少:

[0155] 本申请提供的以Strip Head上的I0的读写策略为拆分判断依据, 不需要对I0以涉及到每个Stripe Head进行拆分。最终拆分为以多个Stripe Head的组合上的子I0。

[0156] (2) I0合并的次数较少:

[0157] 由于拆分的I0数较少, I0最终读写盘时, 做I0合并的计算减少。

[0158] (3) 读写盘的次数较少:

[0159] 本申请是以Stripe Head的组合上的子I0进行计算, 为了计算校验盘或者恢复故障盘上损坏的数据不需要单独每次读写Stripe Head上的数据。

[0160] 通过上述实施例, 可以解决现有技术中, 对读写指令的拆分次数较多, 导致读写指令的执行效率降低, 响应读写指令的延时增大, IOPS降低的问题, 进而达到提升读写指令的执行效率的效果。

[0161] 在本实施例中还提供了一种读写指令的划分装置, 该装置用于实现上述实施例及优选实施方式, 已经进行过说明的不再赘述。如以下所使用的, 术语“模块”可以实现预定功能的软件和/或硬件的组合。尽管以下实施例所描述的装置较佳地以软件来实现, 但是硬件, 或者软件和硬件的组合的实现也是可能并被构想的。

[0162] 图13是根据本申请实施例的读写指令的划分装置的结构框图, 如图13所示, 该装置包括:

[0163] 第一确定模块132, 用于确定接收到的读写指令对目标磁盘的读写信息, 其中, 所

述读写信息用于对所述目标磁盘的数据盘进行读写；

[0164] 第二确定模块134,用于根据所述读写信息确定用于对第一条带头上的第一数据盘进行读写的第一读写指令,以及对第二条带头上的第二数据盘进行读写的第二读写指令,其中,所述第一条带头和所述第二条带头为所述目标磁盘中同一条带中的相邻的两个条带头,所述读写指令至少包括:所述第一读写指令和所述第二读写指令;

[0165] 第三确定模块136,用于确定所述第一读写指令对应的第一读写策略和所述第二读写指令对应的第二读写策略是否一致;

[0166] 划分模块138,用于根据所述第一读写策略和所述第二读写策略是否一致对所述读写指令进行划分。

[0167] 通过上述装置,由于确定接收到的用于对目标磁盘的数据盘进行读写的读写信息;根据读写信息确定第一读写指令以及第二读写指令,其中,第一读写指令用于对第一条带头上的第一数据盘进行读写,第二读写指令用于对第二条带头上的第二数据盘进行读写,第一条带头和第二条带头为目标磁盘中同一条带中的相邻的两个条带头;根据判断第一读写指令对应的第一读写策略与第二读写指令对应的第二读写策略是否一致划分读写指令。因此,可以解决现有技术中,对读写指令的拆分次数较多,导致读写指令的执行效率降低,响应读写指令的延时增大,IOPS降低的问题,进而达到提升读写指令的执行效率的效果。

[0168] 在一个示例性实施例中,所述划分模块,还用于在所述第一读写策略和所述第二读写策略一致的情况下,将所述第一读写指令和所述第二读写指令划分为同一子读写指令;在所述第一读写策略和所述第二读写策略不一致的情况下,将所述第一读写指令和所述第二读写指令划分为不同的子读写指令。

[0169] 在一个示例性实施例中,所述装置还包括:第四确定模块,用于:确定所述读写指令对应的子读写指令集合;确定所述子读写指令集合中的每一子读写指令对应的第三读写策略。

[0170] 在一个示例性实施例中,所述第四确定模块,还用于确定所有所述子读写指令对应的第三数据盘是否为故障盘;根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略。

[0171] 在一个示例性实施例中,所述第四确定模块,还用于在所有所述第三数据盘均不为故障盘的情况下,确定所述第三读写策略为直接读策略;在任一所述第三数据盘为故障盘的情况下,确定所述第三读写策略为重构读策略。

[0172] 在一个示例性实施例中,所述装置还包括:第五确定模块,用于确定所述子读写指令集合中的子读写指令的数量;确定所述子读写指令的数量与预设数量的第一大小关系;根据所述第一大小关系确定所述第三读写策略对应的执行方式。

[0173] 在一个示例性实施例中,上所述第五确定模块,还用于在所述第一大小关系指示所述数量等于所述预设数量的情况下,直接执行所述第三读写策略;在所述第一大小关系指示所述数量大于所述预设数量的情况下,获取所述条带在磁盘中的第一源地址,并将所述第一源地址存储至所述磁盘对应的芯片上的第一存储空间;根据所述第一源地址确定每个所述子读写指令对应的第二源地址,将每个所述第二源地址存储至不同的第二存储空间;根据所述第二存储空间中的第二源地址执行所述第三读写策略。

[0174] 在一个示例性实施例中,所述第五确定模块,还用于根据所述第二源地址确定所述子读写指令对应的第三条带头;响应所述子读写指令,以在所述第三条带头中读取所述第三数据盘中的数据。

[0175] 在一个示例性实施例中,所述第五确定模块,还用于根据所述第二源地址确定所述子读写指令对应的第三条带头;读取所述第三条带头中第四数据盘中的第一数据和所述第三条带头中校验盘的第二数据,其中,所述第四数据盘为所述第三条带头中除所述第三数据盘之外的数据盘;根据所述第一数据和所述第二数据确定所述子读写指令对应的数据盘对应的数据。

[0176] 在一个示例性实施例中,所述第四确定模块,还用于确定所有所述子读写指令对应的第三数据盘是否为故障盘;根据所有所述第三数据盘是否为故障盘确定所述每一子读写指令对应的第三读写策略。

[0177] 在一个示例性实施例中,所述第四确定模块,还用于在所有所述第三数据盘均不为所述故障盘的情况下,根据所述条带中的故障盘的类型确定所述每一子读写指令对应的第三读写策略;在任一所述第三数据盘为所述故障盘的情况下,确定所述第三读写策略为重构写策略。

[0178] 在一个示例性实施例中,所述第四确定模块,还用于在所述条带中的故障盘的类型为数据盘的情况下,确定所述第三读写策略为读修改写策略;在所述条带中的故障盘的类型为校验盘的情况下,确定所述第三读写策略为直接写策略。

[0179] 在一个示例性实施例中,所述装置还包括:第六确定模块,用于确定所述子读写指令集合中的子读写指令的数量;确定所述子读写指令的数量与预设数量的第二大小关系;根据所述第二大小关系确定所述第三读写策略的执行方式。

[0180] 在一个示例性实施例中,所述第六确定模块,还用于在所述第二大小关系指示所述数量等于预设数量的情况下,直接执行所述第三读写策略;在所述第二大小关系指示所述数量大于预设数量的情况下,获取每个所述子读写指令对应的第四条带头中存储的第四数据,并将每个所述第四数据存储至所述磁盘对应的芯片上的不同的第三存储空间;对所述第三存储空间上的第四数据执行所述第三读写策略。

[0181] 在一个示例性实施例中,所述第六确定模块,还用于获取所述第三存储空间中的第五数据,其中,所述第五数据为所述第四条带头中第五数据盘中存储的数据,所述第五数据盘为所述第四条带头中除所述第三数据盘之外的数据盘;获取所述子读写指令中携带的待写入数据;根据所述第五数据和所述待写入数据计算校验数据;将所述第三数据盘对应的数据更新为所述待写入数据,以及将所述第四条带头中校验盘对应的数据更新为所述校验数据。

[0182] 在一个示例性实施例中,所述第六确定模块,还用于获取所述第三存储空间中的第五数据和第六数据,其中,所述第五数据为所述第四条带头中的第五数据盘中存储的数据,所述第五数据盘为所述第四条带头中除所述第三数据盘之外的数据盘,所述第六数据为所述第四条带头中的校验盘中存储的数据;获取所述子读写指令中携带的待写入数据;根据所述第五数据、所述第六数据和所述待写入数据计算校验数据;将所述第三数据盘对应的数据更新为所述待写入数据,以及将所述第六数据更新为所述校验数据。

[0183] 在一个示例性实施例中,所述第六确定模块,还用于获取所述子读写指令中携带

的待写入数据;将所述第三数据盘对应的数据更新为所述待写入数据,并禁止对所述第四条带头中校验盘对应的数据进行更新。

[0184] 在一个示例性实施例中,所述第六确定模块,还用于确定所述第三存储空间上的更新后的第三数据;确定所述第三存储空间对应的条带头;将所述更新后的第三数据存储至所述第三存储空间对应的条带头中。

[0185] 在一个示例性实施例中,所述第一确定模块,还用于根据所述读写信息确定所述读写指令对所述目标磁盘的一个或者多个条带进行读写操作;在确定所述读写指令对多个条带进行读写操作的情况下,将所述读写指令划分为多个目标读写指令,其中,任一所述目标读写指令用于对任一条带进行读写操作,任一所述目标读写指令包括:所述第一读写指令和所述第二读写指令。

[0186] 需要说明的是,上述各个模块是可以通过软件或硬件来实现的,对于后者,可以通过以下方式实现,但不限于此:上述模块均位于同一处理器中;或者,上述各个模块以任意组合的形式分别位于不同的处理器中。

[0187] 本申请的实施例还提供了一种计算机可读存储介质,该计算机可读存储介质中存储有计算机程序,其中,该计算机程序被设置为运行时执行上述任一项方法实施例中的步骤。

[0188] 在一个示例性实施例中,上述计算机可读存储介质可以包括但不限于:U盘、只读存储器(Read-Only Memory,简称为ROM)、随机存取存储器(Random Access Memory,简称为RAM)、移动硬盘、磁碟或者光盘等各种可以存储计算机程序的介质。

[0189] 本申请的实施例还提供了一种电子设备,包括存储器和处理器,该存储器中存储有计算机程序,该处理器被设置为运行计算机程序以执行上述任一项方法实施例中的步骤。

[0190] 在一个示例性实施例中,上述电子设备还可以包括传输设备以及输入输出设备,其中,该传输设备和上述处理器连接,该输入输出设备和上述处理器连接。

[0191] 本实施例中的具体示例可以参考上述实施例及示例性实施方式中所描述的示例,本实施例在此不再赘述。

[0192] 显然,本领域的技术人员应该明白,上述的本申请各模块或各步骤可以用通用的计算装置来实现,它们可以集中在单个的计算装置上,或者分布在多个计算装置所组成的网络上,它们可以用计算装置可执行的程序代码来实现,从而,可以将它们存储在存储装置中由计算装置来执行,并且在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤,或者将它们分别制作成各个集成电路模块,或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样,本申请不限制于任何特定的硬件和软件结合。

[0193] 以上所述仅为本申请的优选实施例而已,并不用于限制本申请,对于本领域的技术人员来说,本申请可以有各种更改和变化。凡在本申请的原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

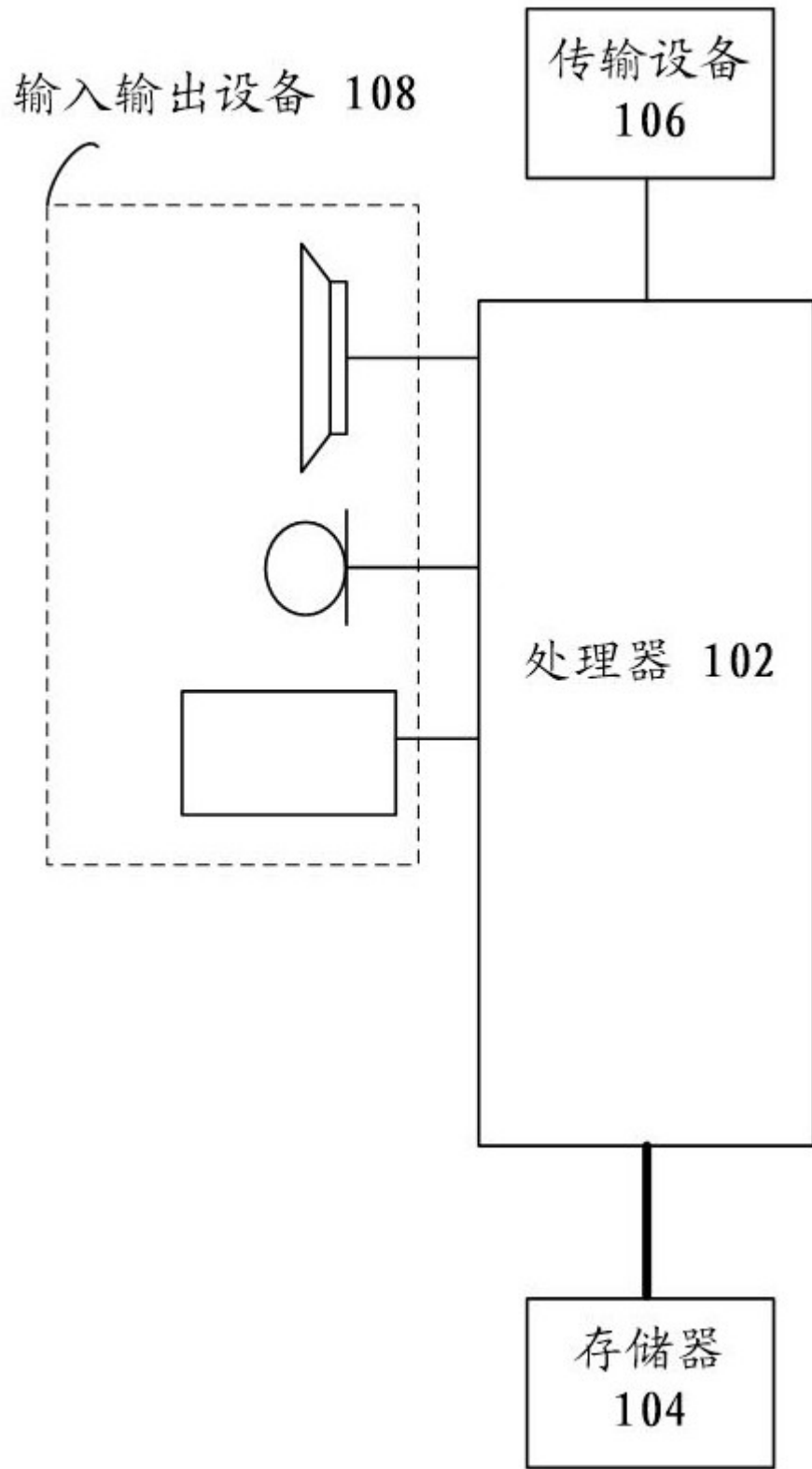


图 1

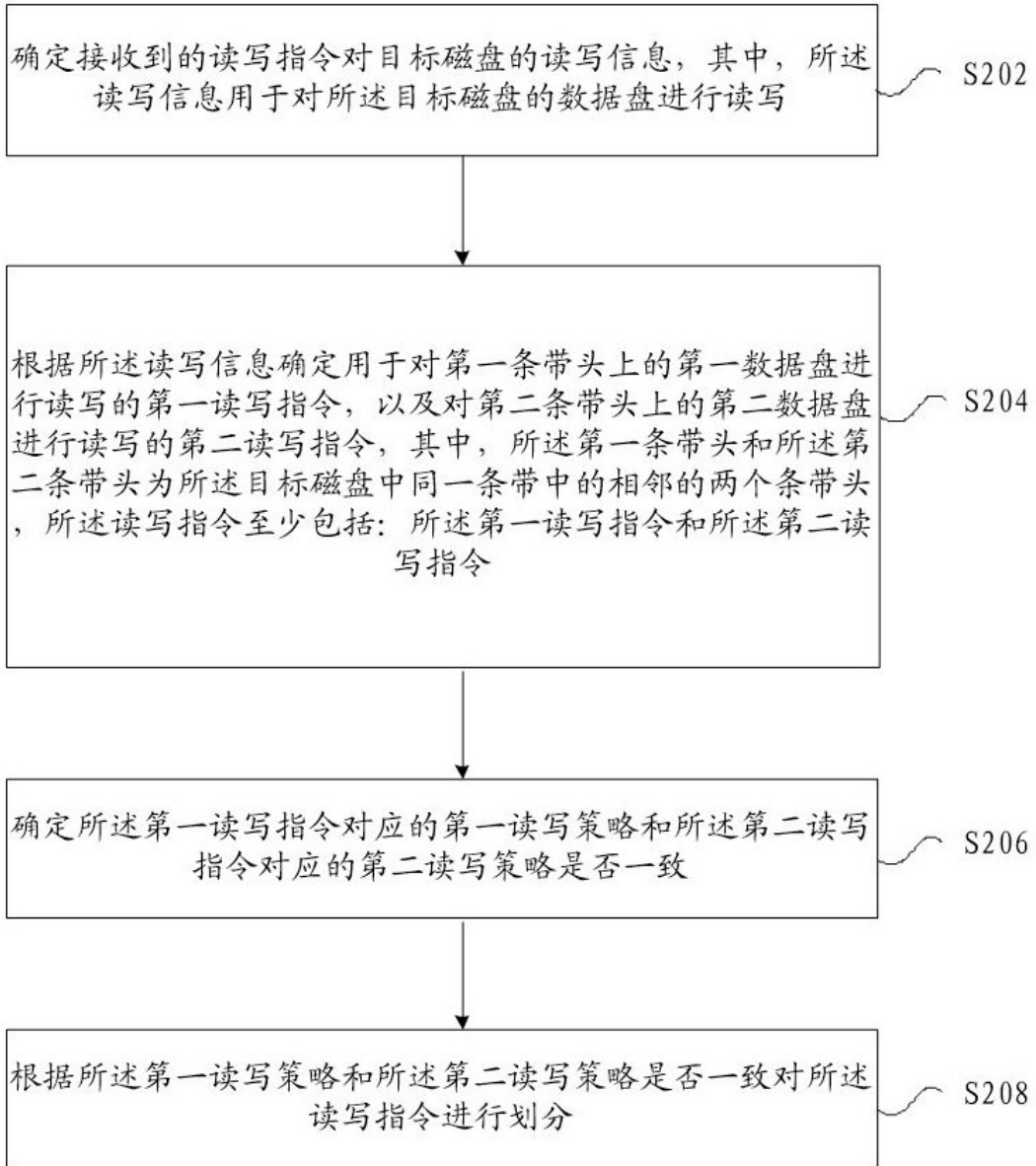


图 2

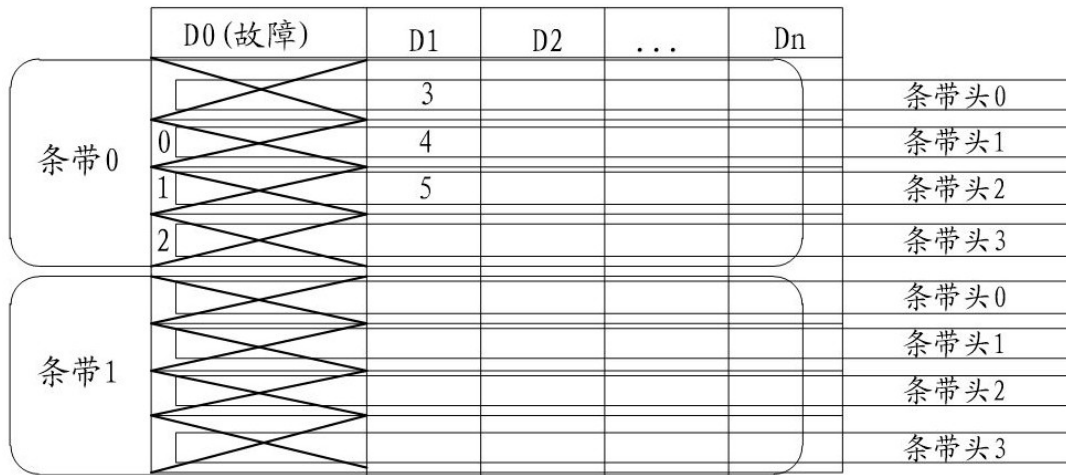


图 3

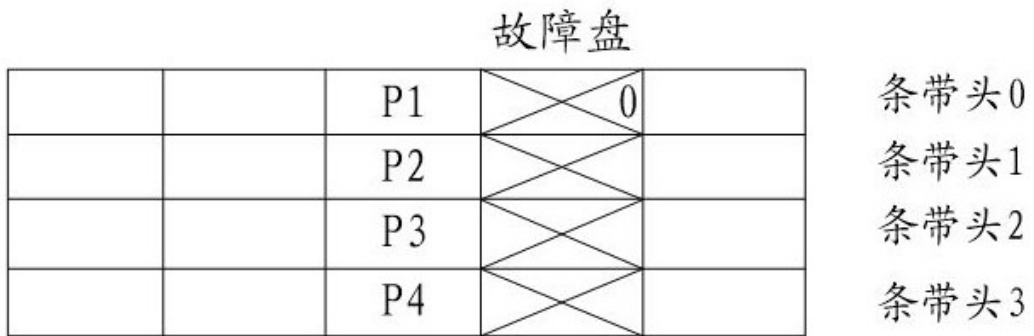


图 4



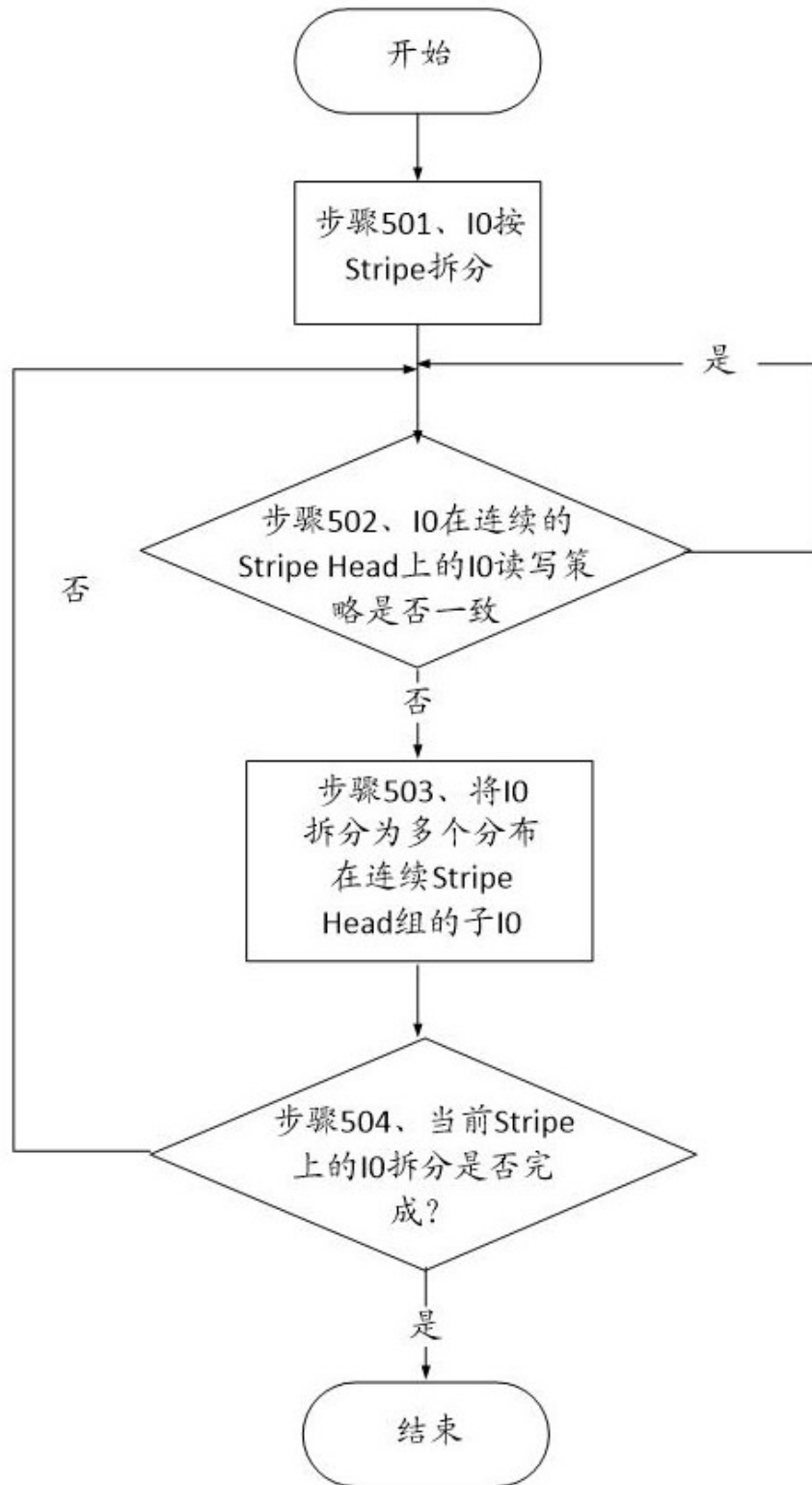


图 5

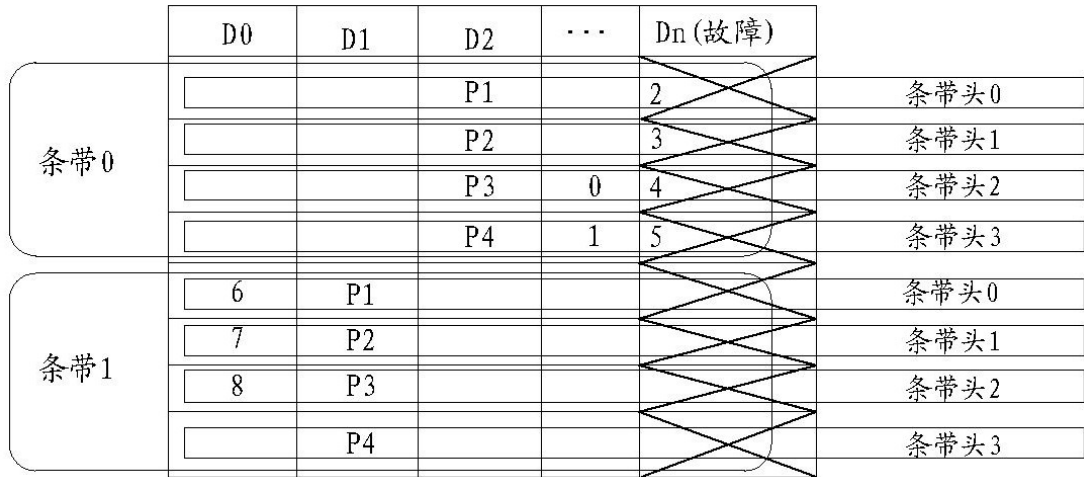


图 6

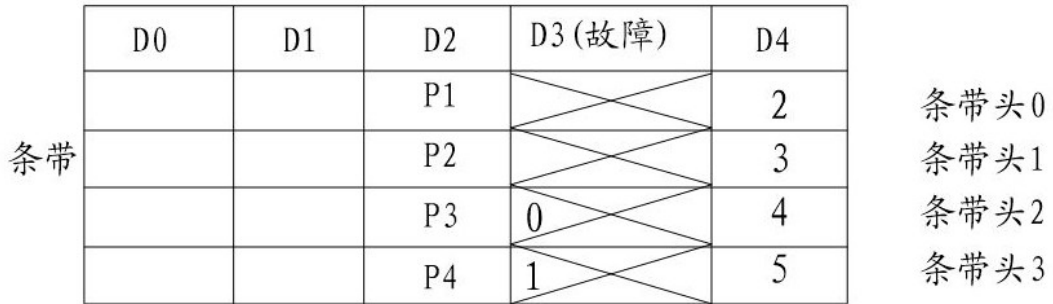


图 7



图 8

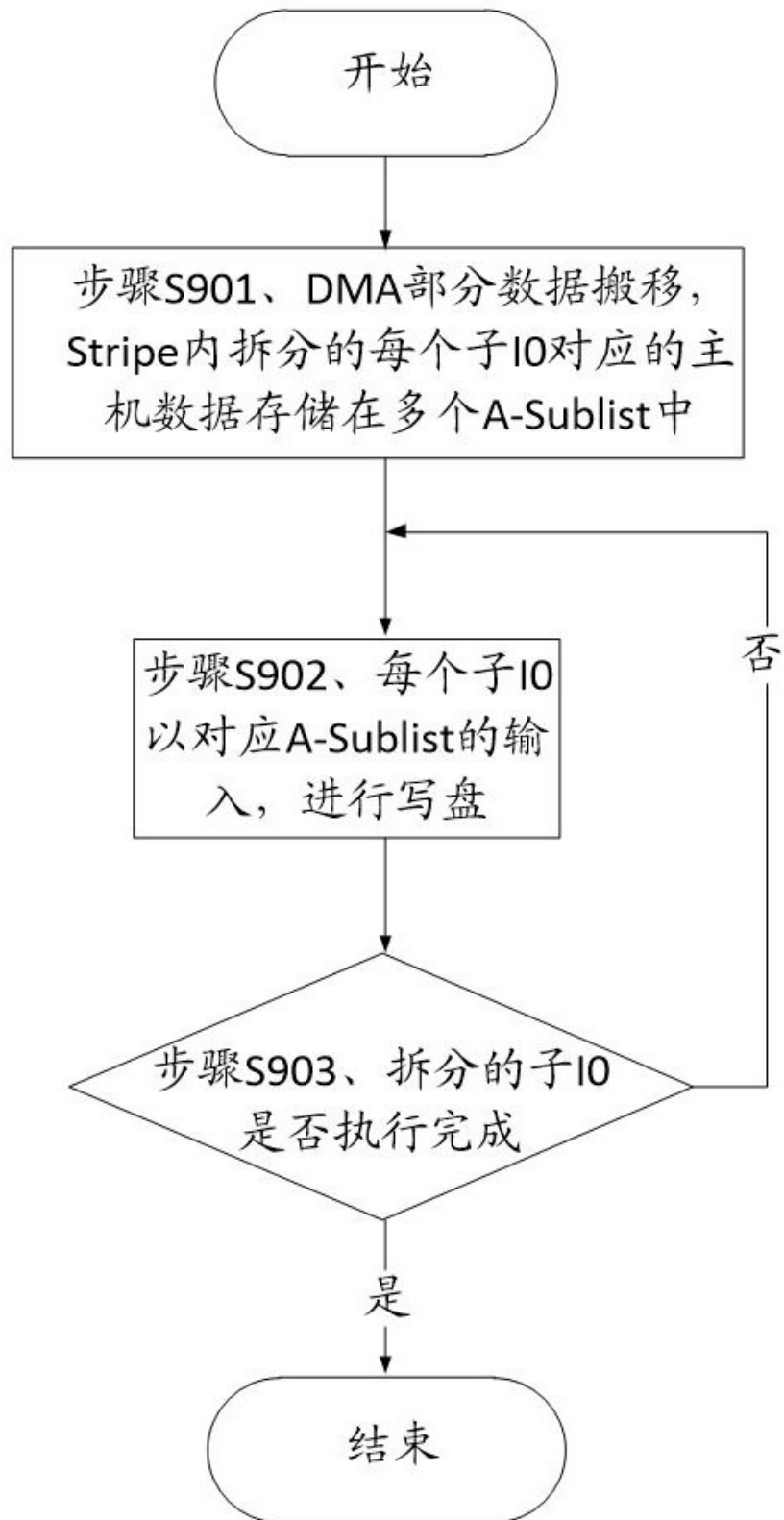


图 9

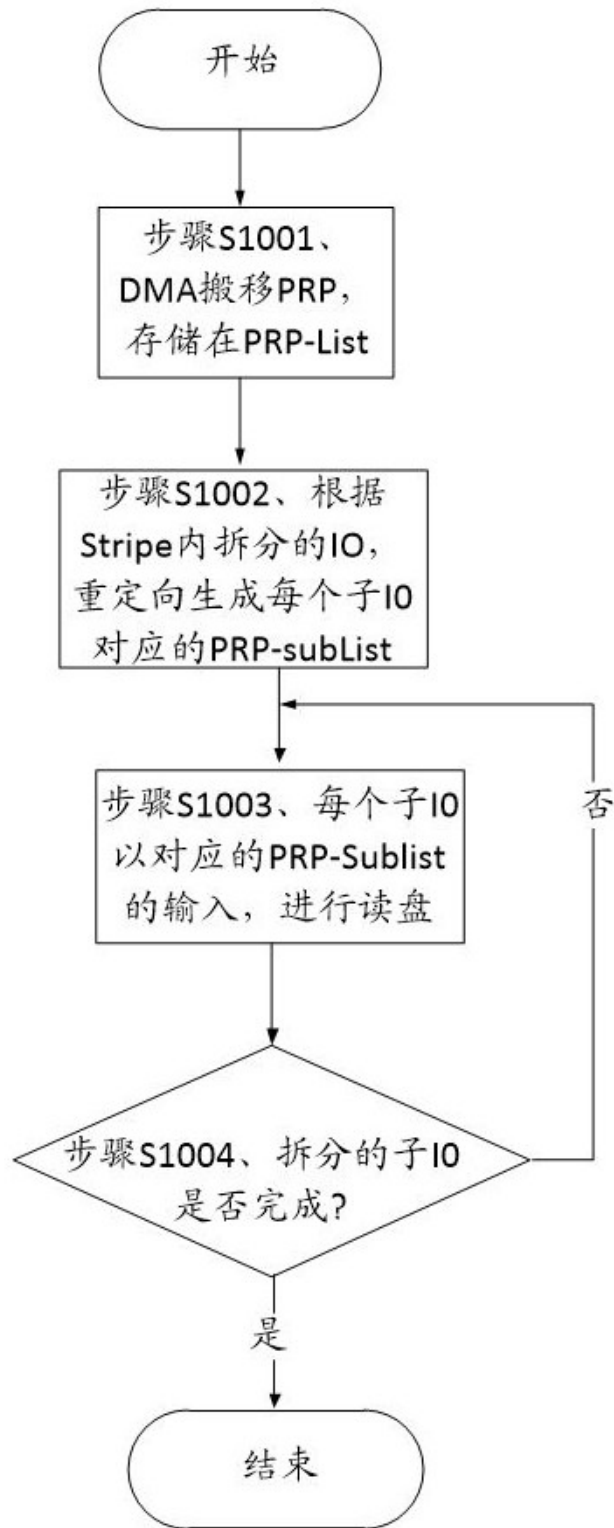


图 10

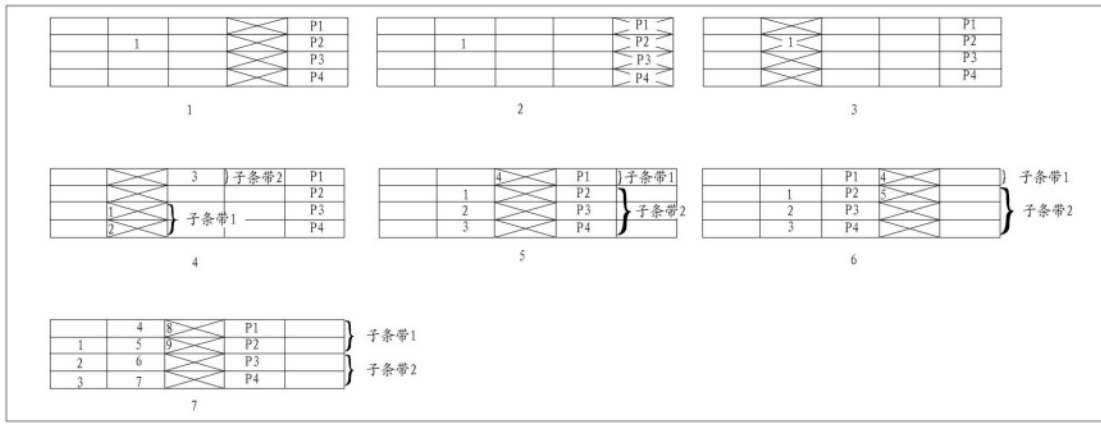


图 11

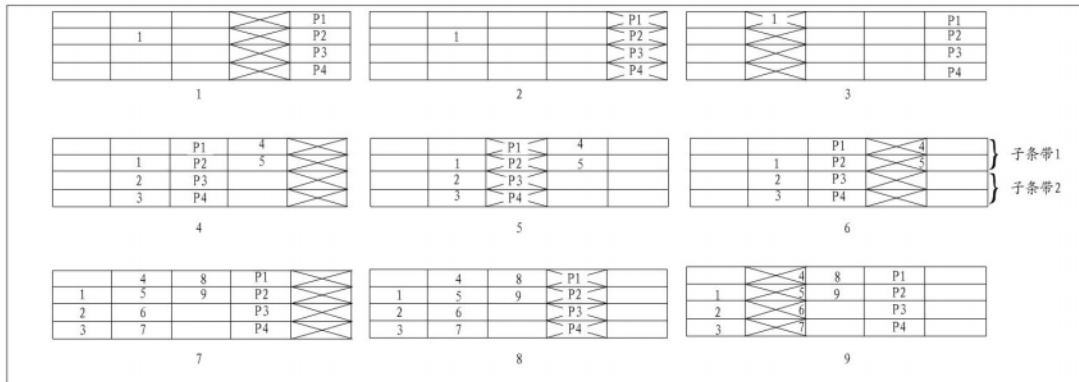


图 12

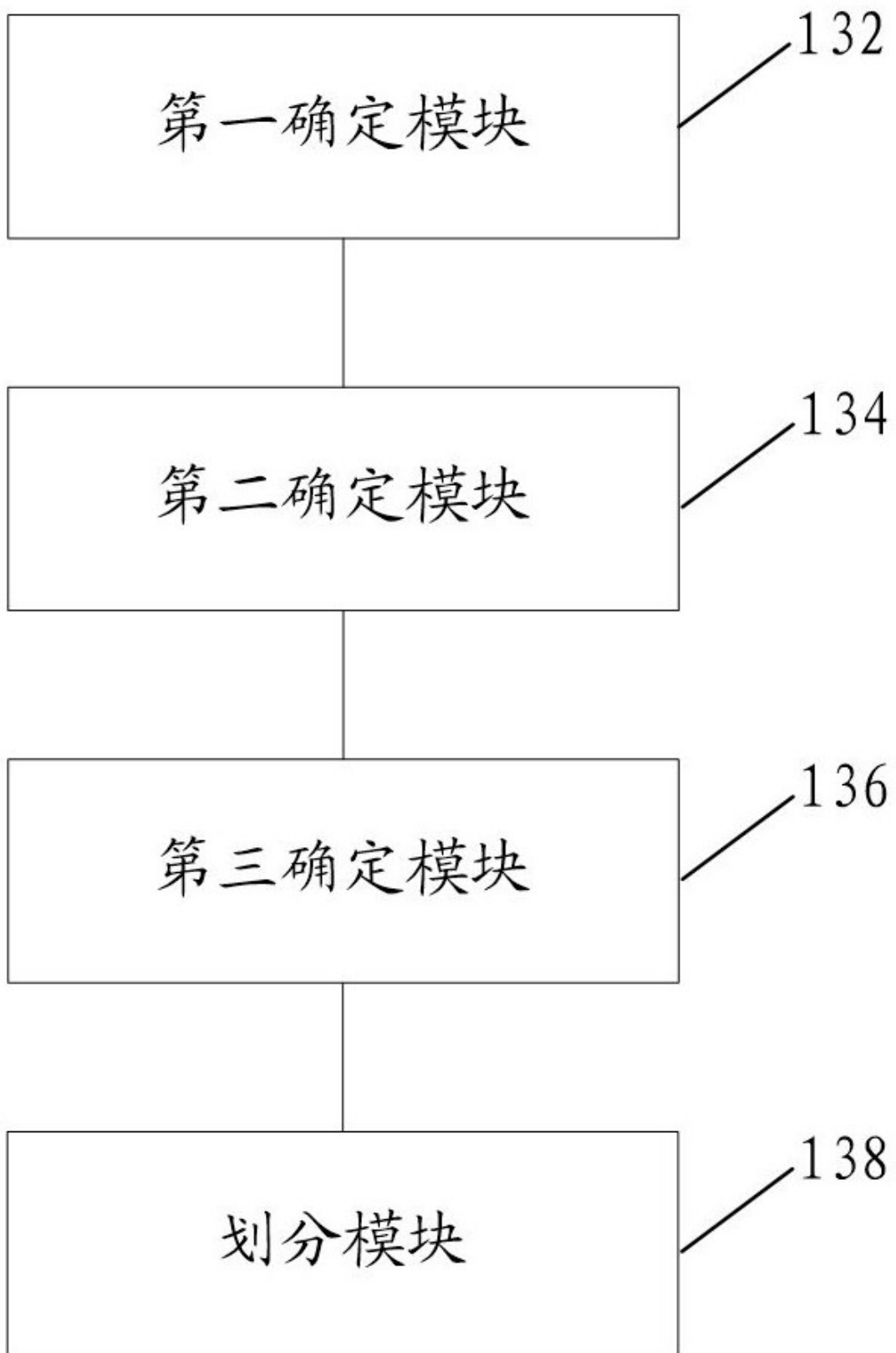


图 13