(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 110659152 B
(45)授权公告日 2021.04.09

(21)申请号 201810691973.2
(22)申请日 2018.06.28
(65)同一申请的已公布的文献号
申请公布号 CN 110659152 A
(43)申请公布日 2020.01.07
(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 李大权 薛强
(74)专利代理机构 北京中博世达专利商标代理有限公司 11274
代理人 申健
(51)Int.Cl.
G06F 11/10(2006.01)

(56)对比文件
CN 102023819 A,2011.04.20
CN 102033716 A,2011.04.27
CN 104461373 A,2015.03.25
US 2012166909 A1,2012.06.28
CN 102023820 A,2011.04.20
CN 102043685 A,2011.05.04

审查员 田翌

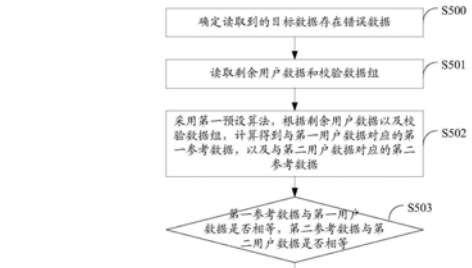
权利要求书2页 说明书12页 附图7页

(54)发明名称

一种数据处理方法及设备

(57)摘要

本申请实施例公开了一种数据处理方法及设备，涉及存储技术领域，解决时延大、计算复杂度高的问题。具体为：确定包括第一用户数据和第二用户数据的目标数据存在错误数据，第一用户数据存储于第一分条的第一数据块，第二用户数据存储于第一分条的第二数据块；读取剩余用户数据和校验数据组，剩余用户数据为第一分条中其他数据块存储的用户数据，校验数据组包括校验块存储的校验数据，第一分条包括至少两个校验块；根据第一预设算法、剩余用户数据和校验数据组，计算与第一用户数据对应的第一参考数据，与第二用户数据对应的第二参考数据；若第一参考数据与第一用户数据相等，且第二参考数据与第二用户数据不等，确定第二用户数据为错误数据。



CN 110659152 B

1.一种数据处理方法,其特征在于,包括:

确定读取到的目标数据存在错误数据,所述目标数据包括第一用户数据和第二用户数据,所述第一用户数据存储于第一数据块中,所述第二用户数据存储于第二数据块中,所述第一数据块和所述第二数据块均属于第一分条;

读取剩余用户数据和校验数据组,所述剩余用户数据包括所述第一分条中除所述第一数据块和所述第二数据块以外的其他数据块中存储的用户数据,所述校验数据组包括所述第一分条中每一校验块存储的校验数据,所述第一分条包括至少两个校验块;

采用第一预设算法,根据所述剩余用户数据以及所述校验数据组,计算得到与所述第一用户数据对应的第一参考数据,以及与所述第二用户数据对应的第二参考数据;

若所述第一参考数据与所述第一用户数据相等,且所述第二参考数据与所述第二用户数据不等,则确定所述第二用户数据为错误数据;

若所述第一参考数据与所述第一用户数据不等,且所述第二参考数据与所述第二用户数据不等,或者,若所述第一参考数据与所述第一用户数据相等,且所述第二参考数据与所述第二用户数据相等,则确定数据丢失。

2.根据权利要求1所述的数据处理方法,其特征在于,所述确定所述第二用户数据为错误数据之后,所述数据处理方法还包括:

采用第二预设算法根据所述第一用户数据、所述剩余用户数据以及所述校验数据组,计算所述第二用户数据的纠正数据;

存储所述第二用户数据的纠正数据,并将已存储的所述第二用户数据置为无效数据。

3.根据权利要求2所述的数据处理方法,其特征在于,

所述确定读取到的目标数据存在错误数据之前,所述数据处理方法还包括:

接收主机发送的读数据请求;

所述计算所述第二用户数据的纠正数据之后,所述数据处理方法还包括:

向所述主机发送所述第一用户数据以及所述第二用户数据的纠正数据。

4.一种存储设备,其特征在于,包括:

索引Index模块,用于确定读取到的目标数据存在错误数据,所述目标数据包括第一用户数据和第二用户数据,所述第一用户数据存储于第一数据块中,所述第二用户数据存储于第二数据块中,所述第一数据块和所述第二数据块均属于第一分条;

独立磁盘冗余阵列RAID模块,用于读取剩余用户数据和校验数据组,所述剩余用户数据包括所述第一分条中除所述第一数据块和所述第二数据块以外的其他数据块中存储的用户数据,所述校验数据组包括所述第一分条中每一校验块存储的校验数据,所述第一分条包括至少两个校验块,以及用于采用第一预设算法,根据所述剩余用户数据以及所述校验数据组,计算得到与所述第一用户数据对应的第一参考数据,以及与所述第二用户数据对应的第二参考数据,以及用于若所述第一参考数据与所述第一用户数据相等,且所述第二参考数据与所述第二用户数据不等,则确定所述第二用户数据为错误数据;

所述RAID模块,还用于若所述第一参考数据与所述第一用户数据不等,且所述第二参考数据与所述第二用户数据不等,或者,若所述第一参考数据与所述第一用户数据相等,且所述第二参考数据与所述第二用户数据相等,则确定数据丢失。

5.根据权利要求4所述的存储设备,其特征在于,

所述RAID模块，还用于在确定所述第二用户数据为错误数据之后，采用第二预设算法根据所述第一用户数据、所述剩余用户数据以及所述校验数据组，计算所述第二用户数据的纠正数据；

所述存储设备还包括存储模块；

所述存储模块，用于存储所述RAID模块计算出的所述第二用户数据的纠正数据，并将已存储的所述第二用户数据置为无效数据。

6.根据权利要求5所述的存储设备，其特征在于，所述存储设备还包括接收模块；

所述接收模块，用于在所述Index模块确定读取到的目标数据存在错误数据之前，接收主机发送的读数据请求；

所述存储设备还包括发送模块；

所述发送模块，用于在所述RAID模块计算所述第二用户数据的纠正数据之后，向所述主机发送所述第一用户数据以及所述第二用户数据的纠正数据。

一种数据处理方法及设备

技术领域

[0001] 本申请实施例涉及存储技术领域，尤其涉及一种数据处理方法及设备。

背景技术

[0002] 独立磁盘冗余阵列(Redundant Array of Independent disks,RAID)是一种把多个独立的磁盘(disk)组合起来形成的一个磁盘阵列，可以向主机提供比单个磁盘更高的存储性能，并可提供数据备份功能。

[0003] 但是，随着制程的演进，磁盘的可擦写周期数逐渐降低，产生误码的因素也越发增多。在这种情况下，部分输入输出(Input Output,IO)需要经过多次读重试或是采用其它处理措施才能读出正确的数据。此时，读IO的时延较大，且读出正确数据所需的计算复杂度较高。

发明内容

[0004] 本申请实施例提供一种数据处理方法及设备，能够解决读IO时延较大、读取正确数据所需的计算复杂度较高的问题。

[0005] 为达到上述目的，本申请采用如下技术方案：

[0006] 第一方面，提供一种数据处理方法，应用于采用磁盘阵列的存储设备(后续均采用存储设备表示)，且该存储设备中的一个分条包括至少两个校验块，该存储设备在确定读取到的包括第一用户数据和第二用户数据的目标数据存在错误数据后，读取第一分条中除第一数据块和第二数据块以外的其他数据块中存储的用户数据(即剩余用户数据)和第一分条中每一校验块存储的校验数据(即校验数据组)，并采用第一预设算法，根据剩余用户数据以及校验数据组，计算得到与第一用户数据对应的第一参考数据，以及与第二用户数据对应的第二参考数据，这样，若第一参考数据与第一用户数据相等，且第二参考数据与第二用户数据不等，则该存储设备确定第二用户数据为错误数据。

[0007] 本申请中的存储设备在确定出目标数据存在错误数据后，利用分条中的其他用户数据和所有校验数据，识别两个用户数据是否为错误数据。相比于现有技术，本申请中的存储设备可迅速、准确的识别出错误数据，且存储设备一次可识别出两个数据块中的用户数据，有效的降低了计算复杂度。

[0008] 可选的，在本申请的一种可能的实现方式中，若第一参考数据与第一用户数据不等，且第二参考数据与第二用户数据不等，或者，若第一参数数据与第一用户数据相等，且第二参考数据与第二用户数据相等，则确定数据丢失。

[0009] 可选的，在本申请的另一种可能的实现方式中，在存储设备确定第二用户数据为错误数据之后，该存储设备还可以采用第二预设算法根据第一用户数据、剩余用户数据以及校验数据组，计算第二用户数据的纠正数据。进一步地，该存储设备存储第二用户数据的纠正数据，并将已存储的第二用户数据置为无效数据。

[0010] 容易理解的是，在存储设备确定出第二用户数据为错误数据的情况下，存储设备

可直接利用除第二用户数据之外的其他数据恢复第二用户数据，以生成第二用户数据的纠正数据，进而，存储设备存储该第二用户数据的纠正数据，同时，存储设备将已存储的第二用户数据置为无效数据，以减少错误数据的扩散。这里，存储设备将已存储的第二用户数据置为无效数据可以表示存储设备删除已存储的第二用户数据，也可以表示为已存储的第二用户数据添加无效标识。

[0011] 可选的，在本申请的另一种可能的实现方式中，存储设备确定读取到的目标数据存在错误数据之前，还可以接收主机发送的读数据请求。相应的，在计算出第二用户数据的纠正数据之后，该存储设备还向主机发送第一用户数据以及第二用户数据的纠正数据。

[0012] 本申请提供的数据处理方法可以应用于主机与存储设备的通信过程中，如主机从存储设备中读取数据的过程；也可以应用于存储设备定时/周期检测分条的一致性的过程中。

[0013] 第二方面，提供一种存储设备，该存储设备包括索引Index模块和RAID模块。具体的，上述Index模块，用于确定读取到的目标数据存在错误数据，该目标数据包括第一用户数据和第二用户数据，第一用户数据存储于第一数据块中，第二用户数据存储于第二数据块中，第一数据块和第二数据块均属于第一分条。上述RAID模块，用于读取剩余用户数据和校验数据组，剩余用户数据包括第一分条中除第一数据块和第二数据块以外的其他数据块中存储的用户数据，校验数据组包括第一分条中每一校验块存储的校验数据，第一分条包括至少两个校验块，以及用于采用第一预设算法，根据剩余用户数据以及校验数据组，计算得到与第一用户数据对应的第一参考数据，以及与第二用户数据对应的第二参考数据，以及用于若第一参考数据与第一用户数据相等，且第二参考数据与第二用户数据不等，则确定第二用户数据为错误数据。

[0014] 进一步地，在本申请的一种可能的实现方式中，上述RAID模块，还用于若第一参考数据与第一用户数据不等，且第二参考数据与第二用户数据不等，或者，若第一参数数据与第一用户数据相等，且第二参考数据与第二用户数据相等，则确定数据丢失。

[0015] 进一步地，在本申请的另一种可能的实现方式中，上述RAID模块，还用于在确定第二用户数据为错误数据之后，采用第二预设算法根据第一用户数据、剩余用户数据以及校验数据组，计算第二用户数据的纠正数据。相应的，本申请的存储设备还包括存储模块，该存储模块用于存储上述RAID模块计算出的第二用户数据的纠正数据，并将已存储的第二用户数据置为无效数据。

[0016] 进一步地，在本申请的另一种可能的实现方式中，本申请的存储设备还包括接收模块，该接收模块，用于在上述Index模块确定读取到的目标数据存在错误数据之前，接收主机发送的读数据请求。此外，本申请的存储设备还包括发送模块，该发送模块用于在上述RAID模块计算第二用户数据的纠正数据之后，向主机发送第一用户数据以及第二用户数据的纠正数据。

[0017] 第三方面，提供一种存储设备，该存储设备包括：一个或多个处理器、存储介质和通信接口。所述存储介质、所述通信接口与所述一个或多个处理器连接；所述存储设备通过所述通信接口与其他设备通信，所述存储器用于存储计算机程序代码，所述计算机程序代码包括指令，当所述一个或多个处理器执行所述指令时，所述存储设备执行如上述第一方面及其各种可能的实现方式的数据处理方法。

[0018] 第四方面，还提供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指令；当其在上述第二方面或上述第三方面所述的存储设备上运行时，使得该存储设备执行如上述第一方面及其各种可能的实现方式的数据处理方法。

[0019] 第五方面，还提供一种包含指令的计算机程序产品，当其在上述第二方面或上述第三方面所述的存储设备上运行时，使得该存储设备执行如上述第一方面及其各种可能的实现方式的数据处理方法。

[0020] 在本申请中，上述存储设备的名字对设备或功能模块本身不构成限定，在实际实现中，这些设备或功能模块可以以其他名称出现。只要各个设备或功能模块的功能和本申请类似，属于本申请权利要求及其等同技术的范围之内。

[0021] 本申请中第二方面、第三方面、第四方面、第五方面及其各种实现方式的具体描述，可以参考第一方面及其各种实现方式中的详细描述；并且，第二方面、第三方面、第四方面、第五方面及其各种实现方式的有益效果，可以参考第一方面及其各种实现方式中的有益效果分析，此处不再赘述。

附图说明

[0022] 图1为本申请实施例中不同级别RAID的结构示意图；

[0023] 图2为本申请实施例中存储设备的硬件结构示意图；

[0024] 图3为本申请实施例中通信系统的结构示意图一；

[0025] 图4为本申请实施例中通信系统的结构示意图二；

[0026] 图5为本申请实施例提供的数据处理方法的流程示意图一；

[0027] 图6为本申请实施例提供的数据处理方法的流程示意图二；

[0028] 图7为本申请实施例提供的数据处理方法的流程示意图三；

[0029] 图8为本申请实施例提供的存储设备的结构示意图一；

[0030] 图9为本申请实施例提供的存储设备的结构示意图二。

具体实施方式

[0031] 在本申请实施例中，“示例性的”或者“例如”等词用于表示作例子、例证或说明。本申请实施例中被描述为“示例性的”或者“例如”的任何实施例或设计方案不应被解释为比其它实施例或设计方案更优选或更具优势。确切而言，使用“示例性的”或者“例如”等词旨在以具体方式呈现相关概念。

[0032] 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”和“第三”等是用于区别不同对象，而不是用于限定特定顺序。

[0033] RAID是一种多个独立的磁盘组合起来形成的一个磁盘阵列，可以向主机提供单个磁盘更高的存储性能。磁盘阵列中的磁盘可以是机械硬盘，也可以是固态硬盘(Solid State Disk,SSD)。

[0034] 在RAID中，数据的存放是分条化存放的，这样可以将数据并行写入到多个成员磁盘中，大大提升用户数据的写入速度，并且可以提供存储的冗余性。

[0035] 经过不断的发展，现在已经拥有了从RAID 0到RAID 5的6种标准级别的RAID级别。另外，还有RAID 6、RAID 7、RAID 10(RAID 1与RAID 0的组合)、RAID 01(RAID 0与RAID 1的

组合)、三重奇偶校验RAID(RAID Triple-parity,RAID-TP)等RAID级别。不同RAID级别代表着不同的存储性能、数据安全性和存储成本。

[0036] 示例性的,以3块物理磁盘(磁盘1、磁盘2、磁盘3)组成的RAID 5的阵列为例,图1的(A)示出了该阵列中数据的分布。一个分条(Stripe)包括两个数据块和一个校验块,数据块用于存储用户数据,校验块用于存储该分条中用户数据的校验数据。当然,在其他更高级别的RAID中,一个分条中可能包括多个校验块(如RAID 6、RAID TP)。在图1的(A)中,d0、d1、d2、d3、d4、d5均为用户数据,分别存储于第一数据块、第二数据块、第三数据块、第四数据块、第五数据块、第六数据块,P1为d0和d1的校验数据,P2为d2和d3的校验数据,P3为d4和d5的校验数据。其中,d0⊕d1=P1,d2⊕d3=P2,d4⊕d5=P3,⊕表示异或算法。d0、d1和P1均属于分条0,d2、d3和P2属于分条1,d4、d5和P3属于分条2。当某一分条中任一数据块的用户数据损坏时,可以使用异或算法通过其他块中的数据将其计算出来,也就是说,RAID 5可以容忍一块成员磁盘的损坏而不丢失用户数据。

[0037] RAID 6为带有两种分布存储的奇偶校验码的独立磁盘结构,是在RAID 5的基础上,额外增加了一个校验位(Q)。RAID 6能够从磁盘阵列中的任意两个并发的磁盘故障中完成数据恢复。

[0038] 示例性的,以4块物理磁盘(磁盘1、磁盘2、磁盘3、磁盘4)组成的RAID 6的阵列为例,图1的(B)示出了该阵列中数据的分布。一个分条包括两个数据块和两个校验块。在图1的(B)中,d0、d1、d2、d3、d4、d5均为用户数据,分别存储于第一数据块、第二数据块、第三数据块、第四数据块、第五数据块、第六数据块,P1和Q1均为d0和d1的校验数据,P2和Q2均为d2和d3的校验数据,P3和Q3均为d4和d5的校验数据。其中,d0⊕d1=P1&Q1,d2⊕d3=P2&Q2,d4⊕d5=P3&Q3,⊕表示异或算法。d0、d1、P1和Q1均属于分条0,d2、d3、P2和Q2属于分条1,d4、d5、P3和Q3属于分条2。

[0039] RAID-TP是在RAID6的基础上,额外增加了一个校验位。RAID-TP能够从磁盘阵列中的任意三个并发的磁盘故障中完成数据恢复。

[0040] 需要说明的是,在实际应用中,上述RAID还可以包括更多或者更少的磁盘。

[0041] 综上,当分条中某一数据块中存储的用户数据损坏或丢失时,可采用异或算法,根据该分条中的校验数据以及其他用户数据恢复出损坏或丢失的用户数据。

[0042] 一般的,采用磁盘阵列的存储设备包括存储控制器、磁盘以及其他外设。图2示出了该存储设备的硬件结构。如图2所示,采用磁盘阵列的存储设备2包括存储控制器20、磁盘21以及其他外设22。

[0043] 存储控制器20主要包括通信接口201、处理器(processor)202、内存(cache)203、磁盘接口204以及系统总线205。通信接口201、处理器202、内存203以及磁盘接口204通过系统总线205完成相互间的通信。

[0044] 通信接口201用于存储控制器20与其他设备通信。示例性的,通信接口201可以为串行连接SCSI接口(Serial Attached SCSI,SAS),也可以为串行高级技术附件(Serial Advanced Technology Attachment,SATA),还可以为UFS等,本申请实施例对此不作具体限定。

[0045] 处理器202可能是中央处理器(Central Processing Unit,CPU),或者是特定集成电路(Application Specific Integrated Circuit,ASIC),或者是被配置成实施本申请实

施例的一个或多个集成电路。在本申请实施例中，处理器202可以用于接收来自主机的写数据请求或者读数据请求、处理所述写数据请求或者所述读数据请求、将所述写数据请求或者所述读数据请求发送给磁盘21以及其他操作。

[0046] 处理器202中还可以包括缓存(图2中未示出)，用于存储各种程序指令。例如，缓存中可以包括闪存转换层(Flash Translation Layer，FTL)。处理器202可以通过FTL确定与逻辑块地址(Logical Block Address，LBA)对应的物理块地址(Physics Block Address，PBA)。或者，处理器202也可以通过其他软件模块来实现类似的功能。

[0047] 内存203用于缓存从外设接收的数据或从磁盘21读取的数据。内存203可以是随机存取存储器(Random Access Memory，RAM)、只读存储器(Read-Only Memory，ROM)、闪存(Flash memory)或固态磁盘(Solid State Disk，SSD)等各种可以存储数据的非短暂性的(non-transitory)可读介质，在此不做限定。举例来说，当接收到写数据请求时，可以将所述写数据请求保存在内存203中，由处理器202对其进行处理。另外，在某些应用场景下，内存203也可以置于存储控制器20的外部。

[0048] 可选的，处理器202中的缓存和内存203可以集成设置或者独立设置，本申请实施例对此不做限定。

[0049] 磁盘接口204用于与磁盘21通信。处理器202将多个磁盘21的存储空间视为统一空间，并管理该统一空间。该统一空间的结构可以参考上述图1中的(A)或图1中的(B)，此处不再进行详细赘述。

[0050] 其他外设22包括动态随机存储器(Dynamic Random Access Memory，DRAM)、电源等支持系统运行的辅助器件。

[0051] 结合图2，如图3所示，上述存储设备2可与主机3通信，存储设备2能够响应主机3下发的输入输出(Input/Output，IO)请求，例如：返回读IO请求所请求的待访问数据、写入写IO请求中待写入的数据。主机3可以是客户端设备，如计算机，也可以是服务器。这里，存储设备2和主机3可以独立设备，也可以集成于同一设备，本申请对此不作具体限定。为了便于理解，本申请主要以存储设备2和主机3独立设置为例进行说明。存储设备2和主机3之间可以采用有线方式连接，也可以采用无线方式连接以形成远程存储的系统。

[0052] 结合上述图3，如图4所示，在软件层面，上述存储设备2包括索引Index模块401、RAID模块402、存储模块41、接收模块42和发送模块43。Index模块401用于处理主机3的读写请求，管理上层访问到存储空间的映射，并检测数据的正确性和完整性，还用于在确定出读取到的数据存在错误数据后，进行降级读操作。RAID模块402用于计算分条中校验块存储的校验数据，将读IO请求中指示的信息转换为存储模块41能识别的指令通过相应接口下发到存储模块41，还用于利用校验数据进行数据重构。存储模块41用于存储用户数据和校验数据。接收模块42用于接收主机3发送的读数据请求。发送模块43用于向主机3发送读取到的数据。

[0053] 具体的，图4示出的通信系统中读取数据的流程为：①、主机3下发读数据请求；②、Index模块401确定与该读数据请求中LBA对应的PBA，并根据该PBA向RAID模块402发送读数据请求；③、响应于读数据请求，RAID模块402从存储模块41中读取目标数据；④、RAID模块402向Index模块401发送目标数据；⑤、Index模块401检测目标数据中是否存在错误数据；⑥、当确定目标数据中未存在错误数据时，Index模块401通过发送模块43向主机3发送该目

标数据。

[0054] 可选的，存储控制器20还可以包括校验模块403。校验模块403用于校验分条的一致性，以保证分条内用户数据的正确性。校验模块403可以周期性的校验分条的一致性，也可以定时校验分条的一致性。

[0055] 具体的，校验分条的一致性的流程为：①、校验模块403向RAID模块402发送读数据请求；②、响应于读数据请求，RAID模块402从磁盘21中读取目标数据；③、RAID模块402向校验模块403发送目标数据；④、校验模块403根据RAID模块发送的目标数据进行分条一致性的校验。

[0056] 上述处理器202用于执行Index模块401、RAID模块402和校验模块403的步骤。

[0057] 现有技术中，当Index模块401检测出目标数据中存在错误数据时，Index模块401向RAID模块402发送降级读信息，以便于RAID模块402从磁盘21中重新读取目标数据，重复执行上述步骤③～⑤，直到确定重新读取到的目标数据不存在错误数据。由于无法确定目标数据中的哪些数据损坏，因此，RAID模块402需要穷举降级组合，计算复杂度为O(n)，n为磁盘的数量，以获取到正确的数据，导致读取到正确数据的时延较大，计算复杂度较高。

[0058] 当校验模块403确定分条不一致后，根据用户配置发告警指示，并认为分条中校验块存储的校验数据出错。这样，无法识别出错误数据，还可能会扩散错误数据。

[0059] 针对上述问题，本申请提供一种数据处理方法，应用于采用磁盘阵列的存储设备(后续均采用存储设备表示)，且该存储设备中的一个分条包括至少两个校验块，该存储设备在确定读取到的包括第一用户数据和第二用户数据的目标数据存在错误数据后，利用目标数据归属的分条中其他数据块所存储的用户数据和该分条中所有校验块中存储的校验数据，计算与第一用户数据对应的第一参考数据以及与第二用户数据对应的第二参考数据，进而，根据第一用户数据与第一参考数据是否相等，以及第二用户数据与第二参考数据是否相等确定出错误数据。相比于现有技术，本申请中的存储设备可迅速、准确的识别出错误数据，且存储设备一次可识别出两个数据块中的用户数据，有效的降低了计算复杂度。

[0060] 本申请实施例提供的数据处理方法适用于在主机向存储设备发送读数据请求后，存储设备的处理过程，也适用于存储设备定期或周期性校验分条的一致性的过程。

[0061] 本申请中存储设备的结构可以参考上述图2或图4，此处不再进行详细赘述。需要说明的是，本申请中存储设备的一个分条包括至少两个校验块，每个校验块均用于存储校验数据。本申请提供的数据处理方法也适用于图3或图4示出的通信系统，此处不再进行详细赘述。

[0062] 以下，结合上述图2示出的存储设备、图3或图4示出的通信系统对本申请提供的数据处理方法进行详细描述。

[0063] 如图5所示，本申请实施例提供的数据处理方法包括：

[0064] S500、存储设备确定读取到的目标数据存在错误数据。

[0065] 这里，目标数据包括第一用户数据和第二用户数据，第一用户数据存储于第一数据块中，第二用户数据存储于第二数据块中，第一数据块和第二数据块均属于第一分条。

[0066] 结合上述描述可知，存储设备中的Index模块用于检测读取到的数据的正确性和完整性。本申请实施例中，存储设备中的Index模块检测读取到的目标数据是否存在错误数据。

[0067] 当Index模块确定目标数据存在错误数据后，Index模块向RAID模块发送降级读请求，指示RAID模块重新读取数据，以便于RAID模块从磁盘中重新读取数据。

[0068] S501、存储设备读取剩余用户数据和校验数据组。

[0069] 剩余用户数据包括第一分条中除第一数据块和第二数据块以外的其他数据块中存储的用户数据，校验数据组包括第一分条中每一校验块存储的校验数据，第一分条包括至少两个校验块。

[0070] 目标数据中的第一用户数据和第二用户数据均存储于第一分条的数据块中。RAID模块在接收到降级读请求后，从磁盘中读取第一分条中除第一数据块和第二数据块以外的其他数据块中存储的用户数据(即读取剩余用户数据)，并从磁盘中读取第一分条中每个校验块存储的校验数据(即读取校验数据组)。

[0071] S502、存储设备采用第一预设算法，根据剩余用户数据以及校验数据组，计算得到与第一用户数据对应的第一参考数据，以及与第二用户数据对应的第二参考数据。

[0072] 具体的，存储设备中的RAID模块采用第一预设算法，根据剩余用户数据以及校验数据组，计算得到与第一用户数据对应的第一参考数据，以及与第二用户数据对应的第二参考数据。

[0073] 可选的，第一预设算法可以为异或算法，也可以为乘加算法，还可以为其他用于解析数据的算法，本申请实施例对此不作具体限定。

[0074] 示例性的，若存储设备包括6块磁盘，存储的数据分别为D0、D1、D2、D3、P、Q，D0、D1、D2以及D3均为用户数据，P和Q均为校验数据，encode(D0,D1,D2,D3)＝(P,Q)，encode()表示某一用于计算校验数据的算法，后续涉及到的encode()与该encode()表示的含义相同，不再一一赘述。若第一用户数据为D0，第二用户数据为D1，D0′表示第一参考数据，D1′表示第二参考数据，则RAID模块可采用公式(D0′,D1′)＝decode(D2,D3,P,Q)计算出D0′和D1′，其中，decode()表示第一预设算法。

[0075] S503、存储设备判断第一参考数据与第一用户数据是否相等，第二参考数据与第二用户数据是否相等。

[0076] 若第一参考数据与第一用户数据相等，且第二参考数据与第二用户数据不等，则存储设备确定第二用户数据为错误数据，即存储设备执行S504。同理，若第一参考数据与第一用户数据不等，且第二参考数据与第二用户数据相等，则存储设备确定第一用户数据为错误数据，即存储设备执行S505。

[0077] 若第一参考数据与第一用户数据不等，且第二参考数据与第二用户数据不等，或者，若第一参数数据与第一用户数据相等，且第二参考数据与第二用户数据相等，则存储设备确定数据丢失，即存储设备执行S506。

[0078] S504、存储设备确定第二用户数据为错误数据。

[0079] S505、存储设备确定第一用户数据为错误数据。

[0080] S506、存储设备确定数据丢失/错误。

[0081] 由于存储设备在识别错误数据的过程中，一次识别两个数据块中的用户数据，因此，在存储设备确定错误数据的过程中，计算复杂度为O(n/2)，n为磁盘数量。相比于现有技术，本申请实施例提供的数据处理方法有效的降低了存储设备的计算复杂度，提高了存储设备确定错误数据的效率。

[0082] 容易理解的是，存储设备在确定出错误数据后，需要将该错误数据进行数据恢复，以避免错误数据的扩散。

[0083] 具体的，结合上述图5，如图6所示，本申请实施例提供的数据处理方法在S504之后还包括S600和S601，在S505之后还包括S602和S603。现对S600、S601、S602和S603进行描述。

[0084] S600、存储设备采用第二预设算法根据第一用户数据、剩余用户数据以及校验数据组，计算第二用户数据的纠正数据。

[0085] 可选的，第二预设算法可以为异或算法，也可以为除减算法，还可以为其他用于解析数据的算法，本申请实施例对此不作具体限定。

[0086] 示例性的，若第二用户数据的纠正数据采用D1''表示，存储设备根据公式encode(D0,D1'',D2,D3)＝(P,Q)计算D1''。

[0087] S601、存储设备存储第二用户数据的纠正数据，并将已存储的第二用户数据置为无效数据。

[0088] 这里“将已存储的第二用户数据置为无效数据”可以用于表示将已存储的第二用户数据删除，也可以用于表示为已存储的第二用户数据添加无效标识。

[0089] 可选的，存储设备在计算出第二用户数据的纠正数据后，可以将已存储的第二用户数据替换为第二用户数据的纠正数据，也可以重新存储第二用户数据的纠正数据，并为已存储的第二用户数据添加无效标识，本申请实施例对此不作具体限定。

[0090] 示例性的，若该存储设备的磁盘阵列由SSD组成，则存储设备重新存储第二用户数据的纠正数据，并为已存储的第二用户数据添加无效标识。这样，后续存储设备在执行垃圾回收的过程中，删除存在无效标识的第二用户数据。

[0091] S602、存储设备采用第二预设算法根据第二用户数据、剩余用户数据以及校验数据组，计算第一用户数据的纠正数据。

[0092] S603、存储设备存储第一用户数据的纠正数据，并将已存储的第一用户数据置为无效数据。

[0093] S602可以参考上述S600的描述，S603可以参考上述S601的描述，此处不再进行详细赘述。不同的是，S600～S601是存储设备对第二用户数据的相关处理过程，S602～S603是存储设备对第一用户数据的相关处理过程。

[0094] 从上述描述可知，本申请实施例提供的数据处理方法适用于在主机向存储设备发送读数据请求后，存储设备的处理过程，也适用于存储设备定期或周期性校验分条的一致性的过程。

[0095] 若该数据处理方法适用于在主机向存储设备发送读数据请求后，存储设备的处理过程，则存储设备在计算出错误数据的纠正数据后，还会向主机发送其获取到的正确数据。具体的，结合上述图6，如图7所示，本申请实施例中的数据处理方法还包括S700，在S600之后还包括S701，在S602之后还包括S702，在S506之后还包括S703。

[0096] S700、主机向存储设备发送读数据请求。

[0097] S701、存储设备向主机发送第一用户数据以及第二用户数据的纠正数据。

[0098] S702、存储设备向主机发送第一用户数据的纠正数据以及第二用户数据。

[0099] S703、存储设备向主机发送数据错误消息，以通知主机待读取数据出现错误。

[0100] 若该数据处理方法适用于存储设备定期或周期性校验分条的一致性的过程，则存

储设备执行S601或S603后该流程结束。

[0101] 上述描述中，目标数据包括第一用户数据和第二用户数据。若上述目标数据还包括存储于第三数据块的第三用户数据和存储于第四数据块的第四用户数据，且第三数据块和第四数据块也属于第一分条，则存储设备在确定上述第一参考数据与第一用户数据不等，且第二参考数据与第二用户数据不等的情况下，采用第一预设算法，根据第一用户数据、第二用户数据、剩余用户数据以及校验数据组，计算得到与第三用户数据对应的第三参考数据，以及与第四用户数据对应的第四参考数据，进而判断第三参考数据与第三用户数据是否相等，以及第四参考数据与第四用户数据是否相等。

[0102] 与上述第一用户数据或第二用户数据类似，若第三参考数据与第三用户数据相等，且第四参考数据与第四用户数据不等，则存储设备确定第四用户数据为错误数据，并采用第二预设算法根据第一用户数据、第二用户数据、第三用户数据、剩余用户数据以及校验数据组，计算第四用户数据的纠正数据。

[0103] 示例性的，若存储设备包括6块磁盘，存储的数据分别为D0、D1、D2、D3、P、Q，D0、D1、D2以及D3均为用户数据，P和Q均为校验数据，encode(D0,D1,D2,D3)＝(P,Q)，存储设备读取到的目标数据包括D0、D1、D2和D3。在确定目标数据存在错误数据后，存储设备执行下述流程：

[0104] (D0′,D1′)＝decode(D2、D3、P、Q)

[0105] 如果D0＝D0′，且D1＝D1′，则确定目标数据存在错误数据，分条一致，错误数据无法修复；

[0106] 如果D0≠D0′，且D1＝＝D1′，则确定D0为错误数据，其他数据正确，此时，错误数据的识别过程结束；

[0107] 如果D0＝D0′，且D1≠D1′，则确定D1为错误数据，其他数据正确，此时，错误数据的识别过程结束；

[0108] 如果D0≠D0′，且D1≠D1′，则重新识别另外两个用户数据，如D2和D3；

[0109] 如果D0、D1、D2、D3、P和Q均无法确定错误数据，则确定数据丢失或者数据出错，且无法修复。

[0110] 综上，相比于现有技术，本申请中的存储设备可迅速、准确的识别出错误数据，且存储设备一次可识别出两个数据块中的用户数据，有效的降低了计算复杂度。此外，本申请提供的数据处理方法还能有效的避免错误数据的扩散。

[0111] 本申请实施例提供一种存储设备，该存储设备用于执行以上数据处理方法中的存储设备所执行的步骤。本申请实施例提供的存储设备可以包括相应步骤所对应的模块。

[0112] 本申请实施例可以根据上述方法示例对存储设备进行功能模块的划分，例如，可以对应各个功能划分各个功能模块，也可以将两个或两个以上的功能集成在一个处理模块中。上述集成的模块既可以采用硬件的形式实现，也可以采用软件功能模块的形式实现。本申请实施例中对模块的划分是示意性的，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。

[0113] 如上述图4所示，在采用对应各个功能划分各个功能模块的情况下，本申请实施例的存储设备包括Index模块401、RAID模块402、存储模块41、接收模块42和发送模块43。

[0114] 其中，上述Index模块401用于支持该存储设备执行S500、S706，和/或用于本文所

描述的技术的其它过程。

[0115] 上述RAID模块402用于支持该存储设备执行上述实施例中的S501、S502、S503、S504、S505、S506、S600、S602和/或用于本文所描述的技术的其它过程。

[0116] 上述存储模块41用于支持该存储设备执行上述实施例中的S601、S603，和/或用于本文所描述的技术的其它过程。

[0117] 上述接收模块42用于支持该存储设备900执行上述实施例中的S700，和/或用于本文所描述的技术的其它过程。

[0118] 上述发送模块43用于支持该存储设备900执行上述实施例中的S701、S702、S703，和/或用于本文所描述的技术的其它过程。

[0119] 进一步地，上述存储设备还包括校验模块403，该校验模块403用于支持该存储设备执行定时或周期性的校验分条的一致性，和/或用于本文所描述的技术的其它过程。

[0120] 需要说明的是，上述方法实施例涉及的各步骤的所有相关内容均可以援引到对应功能模块的功能描述，在此不再赘述。

[0121] 在采用集成的单元的情况下，本申请实施例提供的存储设备的结构示意图如图8所示。在图8中，存储设备包括：处理模块80、通信模块81和存储模块82。处理模块80用于对存储设备的动作进行控制管理，例如，执行上述Index模块401、RAID模块402、校验模块403执行的步骤，和/或用于执行本文所描述的技术的其它过程。通信模块81用于支持存储设备与其他设备之间的交互，例如，执行上述接收模块42和发送模块43执行的步骤。存储模块82用于存储存储设备的程序代码和数据，例如存储上述存储模块41所保存的内容。

[0122] 上述处理模块80可以是处理器或控制器，例如可以是中央处理器(Central Processing Unit，CPU)，数字信号处理器(Digital Signal Processor，DSP)。其可以实现或执行结合本申请实施例公开内容所描述的各种示例性的逻辑方框，模块和电路。通信模块81可以是收发器、收发电路或收发器等。存储模块82可以是存储器。

[0123] 当处理模块80为处理器，通信模块81为收发器，存储模块82为存储器时，本申请实施例所涉及的存储设备可以为图9所示的存储设备。

[0124] 如图9所示，该存储设备包括：收发器90、处理器91和存储器92。其中，收发器90、处理器91与存储器92之间通过系统总线93连接，并完成相互间通信。

[0125] 当存储设备运行时，该存储设备执行如图5～图7中任一附图所示的实施例的数据处理方法。具体的数据处理方法可参见上述如图5～图7中任一附图所示的实施例中的相关描述，此处不再赘述。

[0126] 其中，收发器90用于与其他设备或通信网络通信，如以太网，WLAN等。

[0127] 其中，存储器92用于存储用户数据和校验数据，还可以用于存储软件程序以及应用模块，处理器91通过运行存储在存储器92的软件程序以及应用模块，从而执行存储设备的各种功能应用以及数据处理。

[0128] 存储器92可主要包括存储程序区920和存储数据区99，其中，存储程序区920可存储操作系统、至少一个功能所需的应用程序，比如计算错误数据的纠正数据等；存储数据区99可存储用户数据和校验数据。

[0129] 其中，存储器92可以是ROM，或可存储静态信息和指令的其他类型的静态存储设备，RAM或者可存储信息和指令的其他类型的动态存储设备，也可以是可擦除可编程只读存

储器(Erasable Programmable ROM,EEPROM)、磁盘存储介质或者其他磁存储设备、或者能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由存储设备存取的任何其他介质,但不限于此。

[0130] 存储器92可以是独立存在,通过系统总线93与处理器91相连接。存储器92也可以和处理器91集成在一起。

[0131] 处理器91是存储设备的控制中心。处理器91利用各种接口和线路连接整个存储设备的各个部分,通过运行或执行存储在存储器92内的软件程序和/或应用模块,以及调用存储在存储器92内的数据,执行存储设备的各种功能和处理数据,从而对存储设备进行整体监控。

[0132] 在具体实现中,作为一种实施例,处理器91可以包括一个或多个CPU,例如图9中的处理器91包括CPU 0和CPU 1。

[0133] 系统总线93可以分为地址总线、数据总线、控制总线等。本申请实施例中为了清楚说明,在图9中将各种总线都示意为系统总线93。

[0134] 结合上述图2,图9中的收发器90与通信接口201对应,处理器91与处理器202对应,存储器92与磁盘21对应。

[0135] 本申请实施例还提供一种计算机可读存储介质,计算机可读存储介质中存储有指令,当其在存储设备上运行时,使得存储设备执行上述数据处理方法实施例所示的方法流程中存储设备执行的各个步骤。

[0136] 在上述实施例中,可以全部或部分的通过软件,硬件,固件或者其任意组合来实现。当使用软件程序实现时,可以全部或部分地以计算机程序产品的形式出现。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行所述计算机程序指令时,全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。所述计算机指令可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存储介质传输,例如,所述计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线(DSL))或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心传输。所述计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。该可用介质可以是磁性介质,(例如,软盘,硬盘、磁带)、光介质(例如,DVD)或者半导体介质(例如固态硬盘Solid State Disk(SSD))等。

[0137] 通过以上的实施方式的描述,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,仅以上述各功能模块的划分进行举例说明,实际应用中,可以根据需要而将上述功能分配由不同的功能模块完成,即将装置的内部结构划分成不同的功能模块,以完成以上描述的全部或者部分功能。

[0138] 在本申请所提供的几个实施例中,应该理解到,所揭露的装置和方法,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,所述模块或单元的划分,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装置,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。

[0139] 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一个物理单元或多个物理单元,即可以位于一个地方,或者也可以分布到多个不同地方。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。

[0140] 另外,在本申请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。

[0141] 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时,可以存储在一个可读取存储介质中。基于这样的理解,本申请实施例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个设备(可以是单片机,芯片等)或处理器(processor)执行本申请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(Read-Only Memory,ROM)、随机存取存储器(Random Access Memory,RAM)、磁碟或者光盘等各种可以存储程序代码的介质。

[0142] 以上所述,仅为本申请的具体实施方式,但本申请的保护范围并不局限于此,任何在本申请揭露的技术范围内的变化或替换,都应涵盖在本申请的保护范围之内。因此,本申请的保护范围应以所述权利要求的保护范围为准。

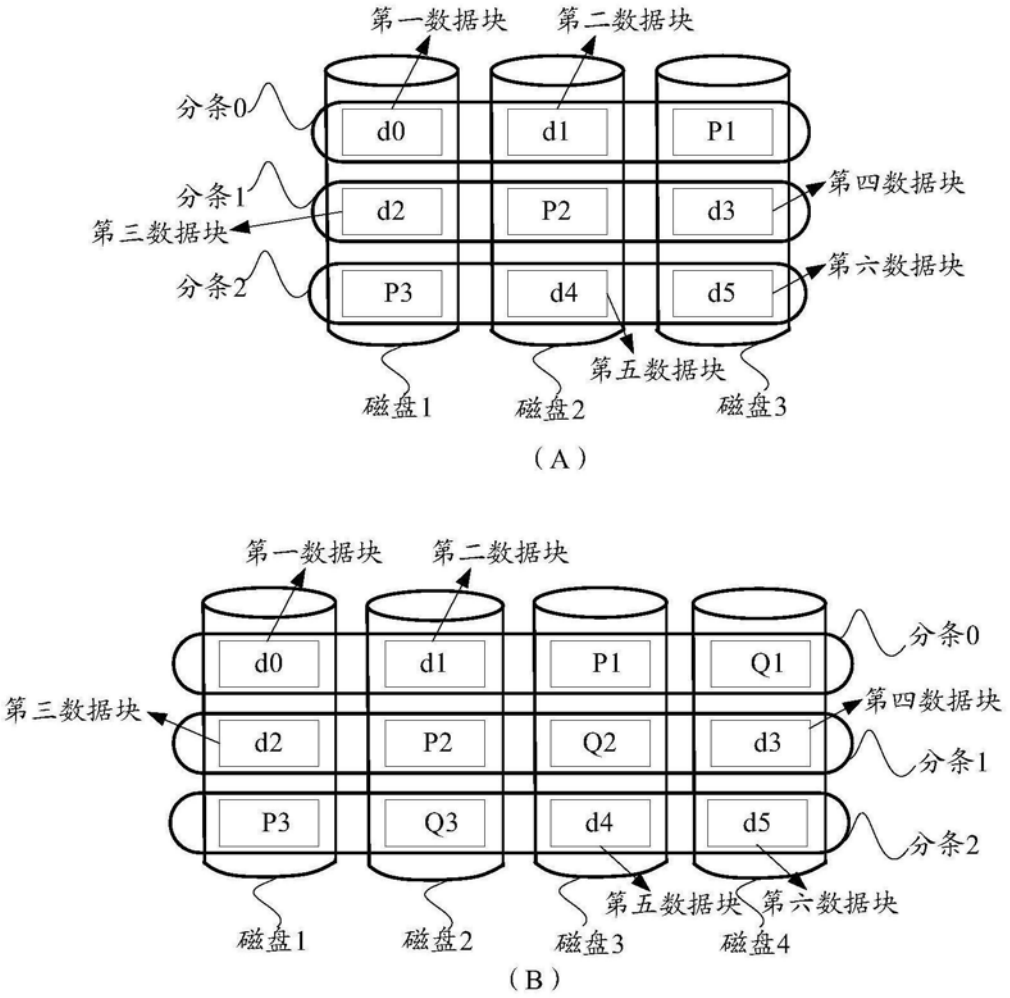

第一数据块
第二数据块
分条0
d0
d1
P1
分条1
d2
P2
d3
第四数据块
第三数据块
分条2
P3
d4
d5
第六数据块
第五数据块
磁盘1
磁盘2
磁盘3
（A）
第一数据块
第二数据块
d0
d1
P1
Q1
分条0
第三数据块
d2
P2
Q2
d3
第四数据块
分条1
P3
Q3
d4
d5
分条2
第五数据块
第六数据块
磁盘1
磁盘2
磁盘3
磁盘4
（B）

图1

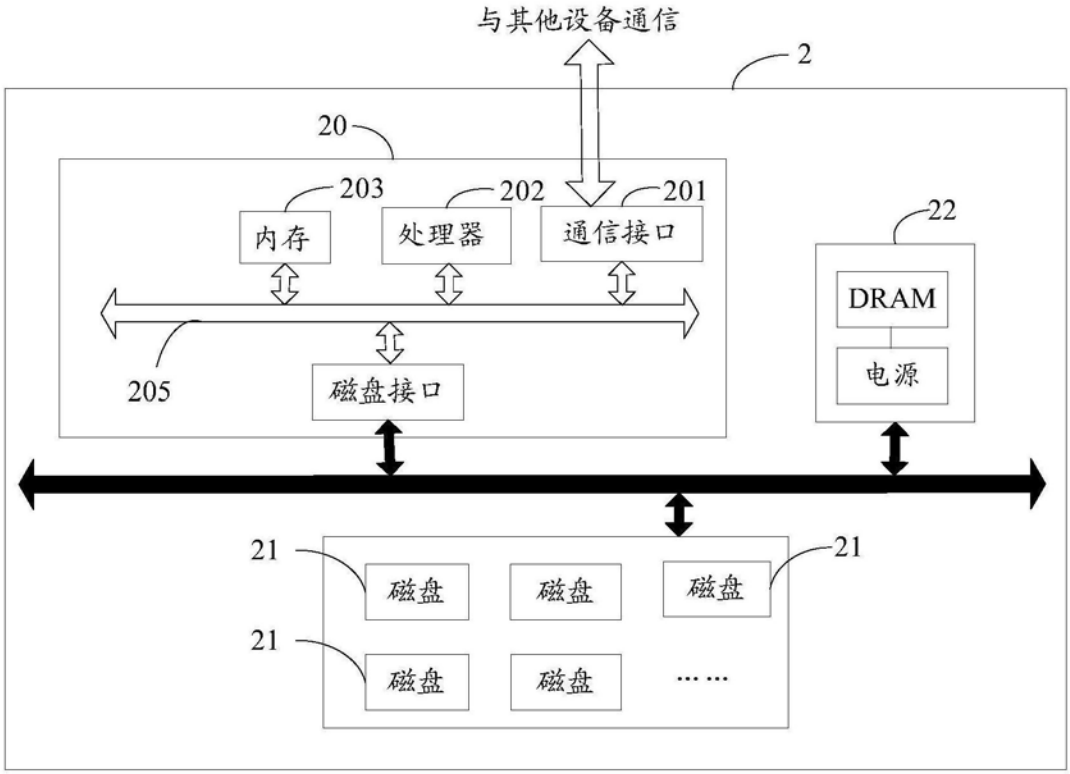

与其他设备通信
2
20
203
202
201
内存
处理器
通信接口
22
DRAM
电源
205
磁盘接口
21
21
磁盘
磁盘
磁盘
21
磁盘
磁盘
……

图2

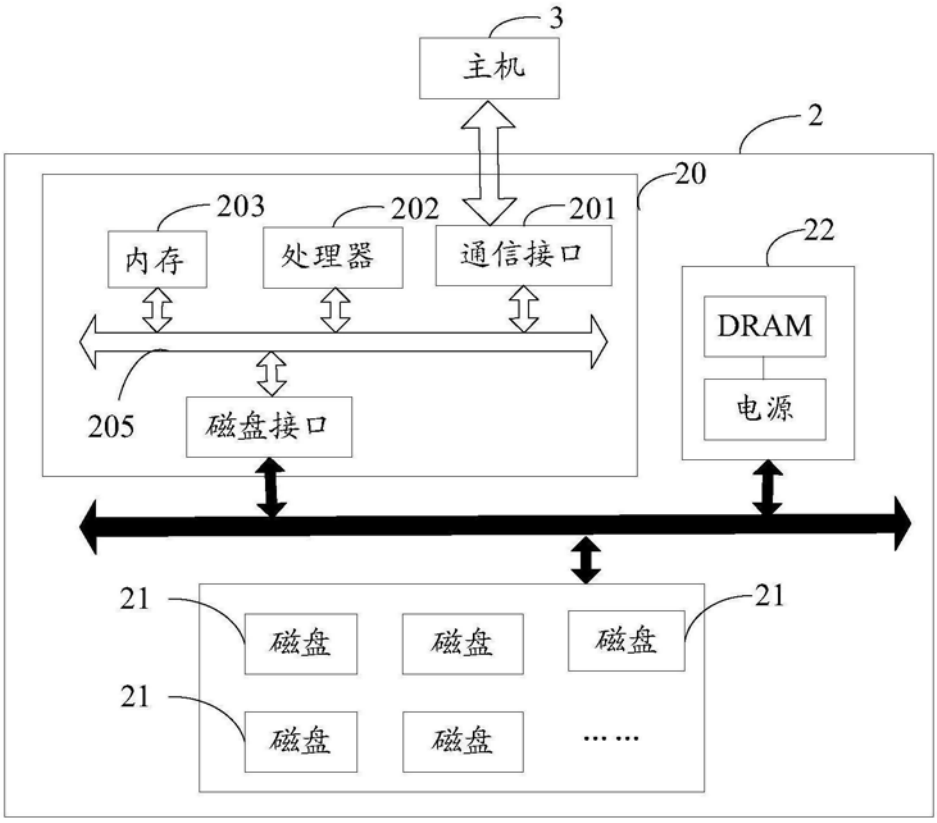

3
主机
2
20
203
202
201
内存
处理器
通信接口
22
DRAM
电源
205
磁盘接口
21
磁盘
磁盘
磁盘
21
21
磁盘
磁盘
……

图3

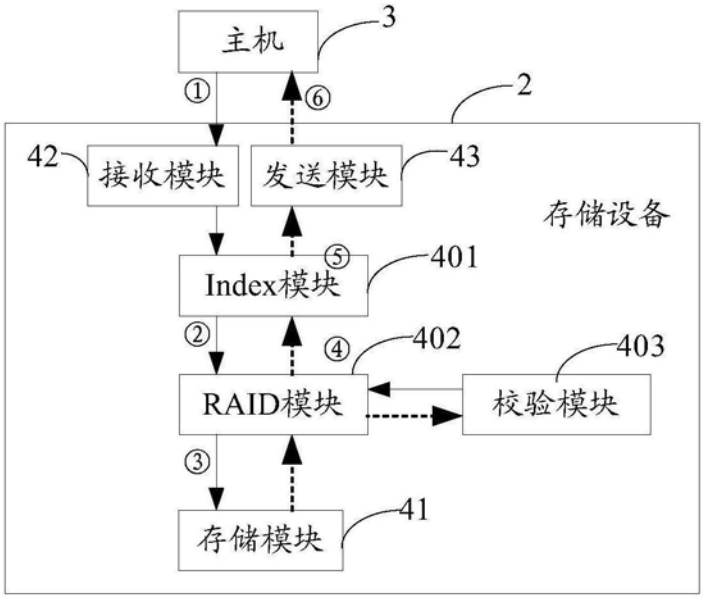

主机
3
①
⑥
2
42
接收模块
发送模块
43
存储设备
⑤
401
Index模块
②
④
402
403
RAID模块
校验模块
③
存储模块
41

图4

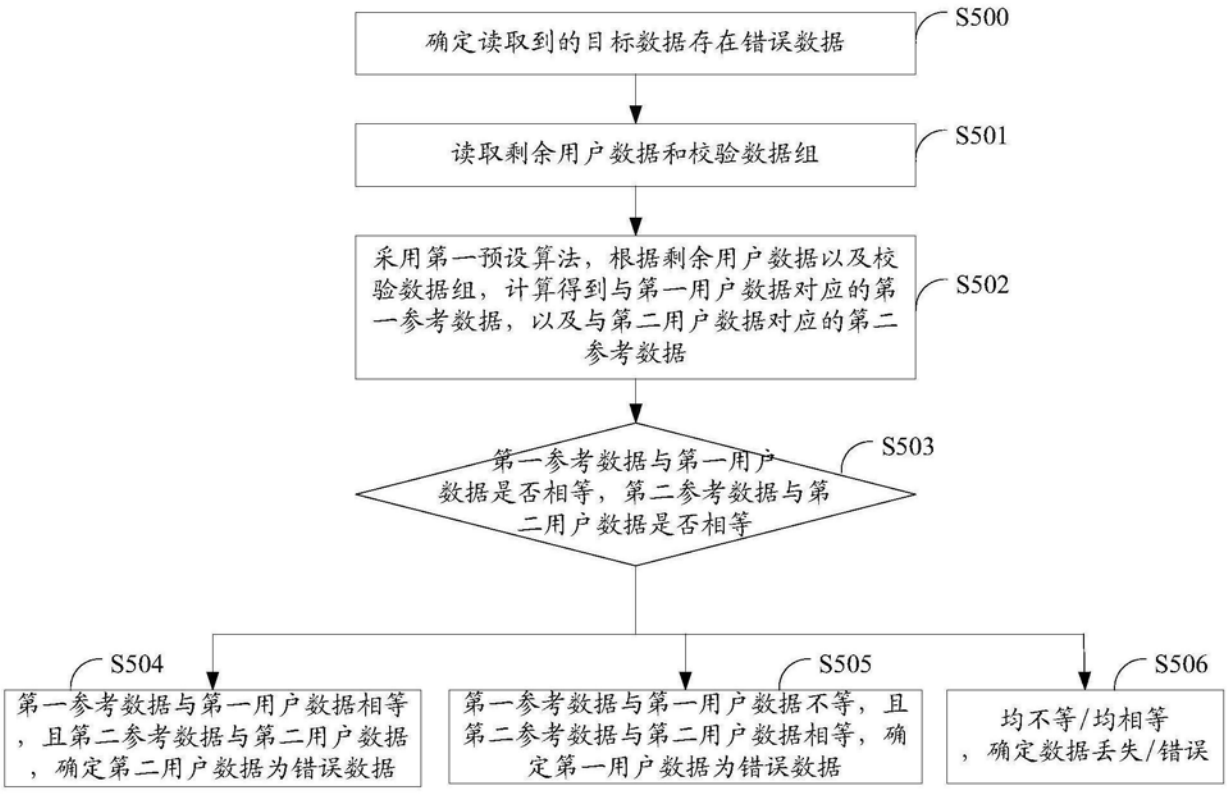

确定读取到的目标数据存在错误数据
S500
读取剩余用户数据和校验数据组
S501
采用第一预设算法，根据剩余用户数据以及校验数据组，计算得到与第一用户数据对应的第一参考数据，以及与第二用户数据对应的第二参考数据
S502
第一参考数据与第一用户数据是否相等，第二参考数据与第二用户数据是否相等
S503
S504
第一参考数据与第一用户数据相等，且第二参考数据与第二用户数据，确定第二用户数据为错误数据
S505
第一参考数据与第一用户数据不等，且第二参考数据与第二用户数据相等，确定第一用户数据为错误数据
S506
均不等/均相等，确定数据丢失/错误

图5

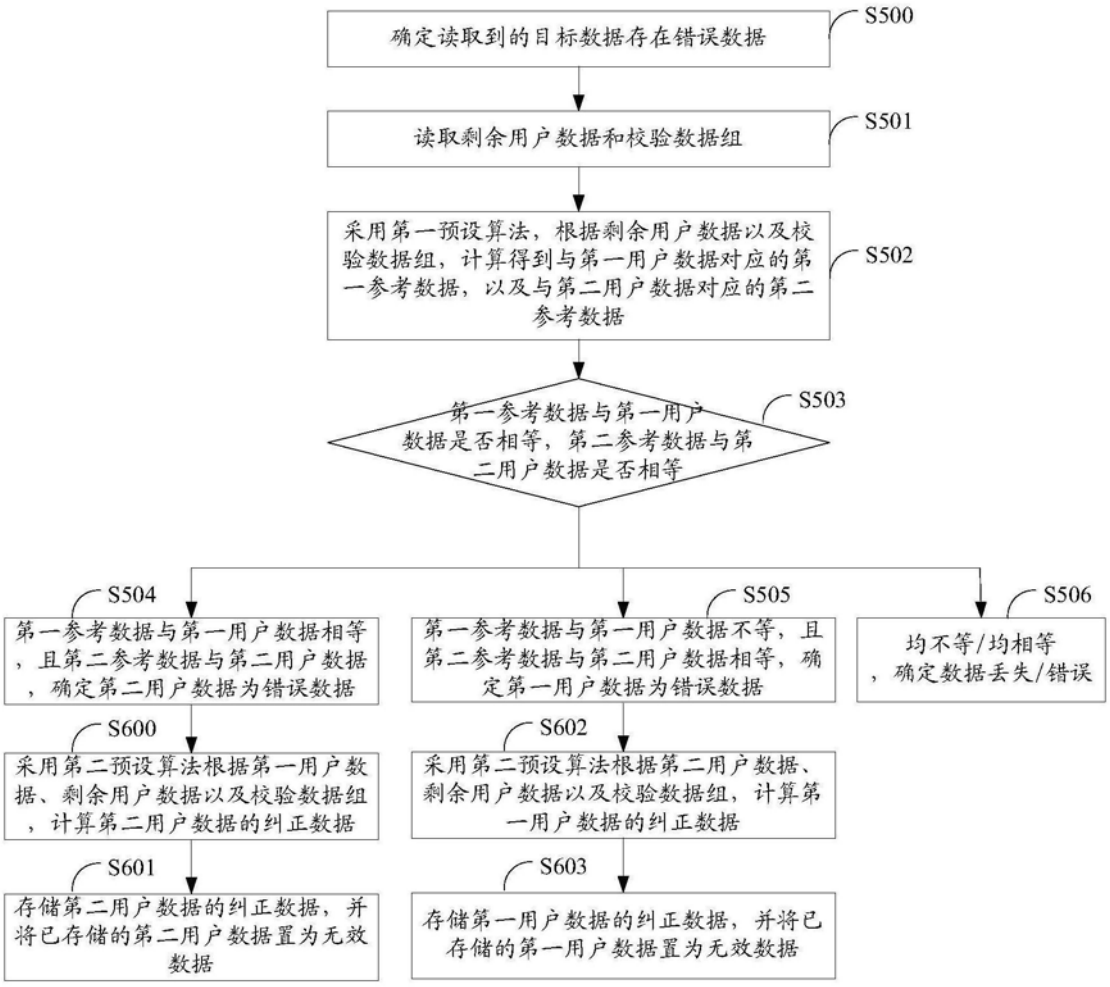

确定读取到的目标数据存在错误数据
S500
读取剩余用户数据和校验数据组
S501
采用第一预设算法，根据剩余用户数据以及校验数据组，计算得到与第一用户数据对应的第一参考数据，以及与第二用户数据对应的第二参考数据
S502
第一参考数据与第一用户数据是否相等，第二参考数据与第二用户数据是否相等
S503
S504
第一参考数据与第一用户数据相等，且第二参考数据与第二用户数据，确定第二用户数据为错误数据
S505
第一参考数据与第一用户数据不等，且第二参考数据与第二用户数据相等，确定第一用户数据为错误数据
S506
均不等/均相等，确定数据丢失/错误
S600
采用第二预设算法根据第一用户数据、剩余用户数据以及校验数据组，计算第二用户数据的纠正数据
S602
采用第二预设算法根据第二用户数据、剩余用户数据以及校验数据组，计算第一用户数据的纠正数据
S601
存储第二用户数据的纠正数据，并将已存储的第二用户数据置为无效数据
S603
存储第一用户数据的纠正数据，并将已存储的第一用户数据置为无效数据

图6

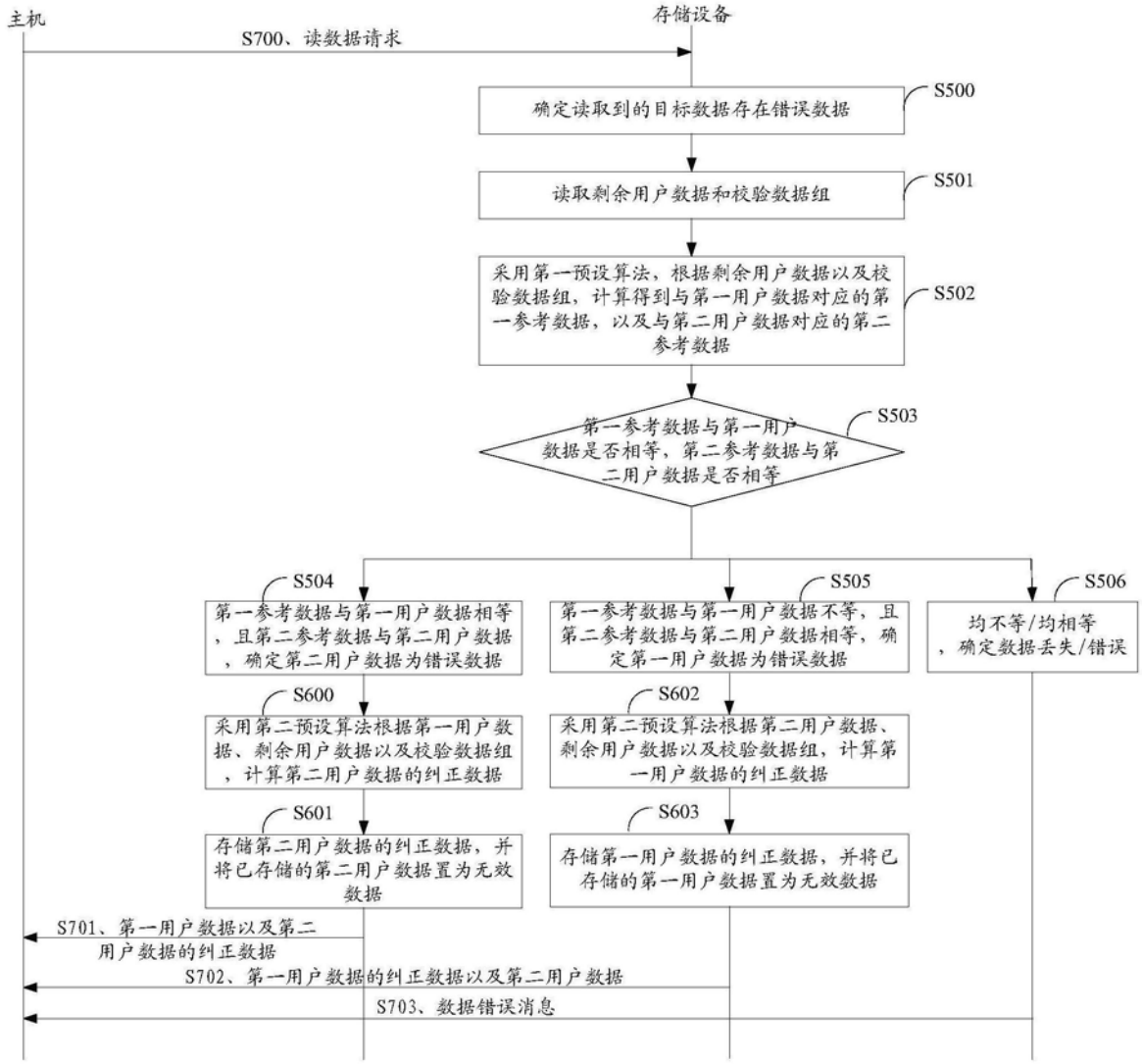

主机
存储设备
S700、读数据请求
确定读取到的目标数据存在错误数据
S500
读取剩余用户数据和校验数据组
S501
采用第一预设算法，根据剩余用户数据以及校验数据组，计算得到与第一用户数据对应的第一参考数据，以及与第二用户数据对应的第二参考数据
S502
第一参考数据与第一用户数据是否相等，第二参考数据与第二用户数据是否相等
S503
S504
第一参考数据与第一用户数据相等，且第二参考数据与第二用户数据，确定第二用户数据为错误数据
S505
第一参考数据与第一用户数据不等，且第二参考数据与第二用户数据相等，确定第一用户数据为错误数据
S506
均不等/均相等，确定数据丢失/错误
S600
采用第二预设算法根据第一用户数据、剩余用户数据以及校验数据组，计算第二用户数据的纠正数据
S602
采用第二预设算法根据第二用户数据、剩余用户数据以及校验数据组，计算第一用户数据的纠正数据
S601
存储第二用户数据的纠正数据，并将已存储的第二用户数据置为无效数据
S603
存储第一用户数据的纠正数据，并将已存储的第一用户数据置为无效数据
S701、第一用户数据以及第二用户数据的纠正数据
S702、第一用户数据的纠正数据以及第二用户数据
S703、数据错误消息

图7

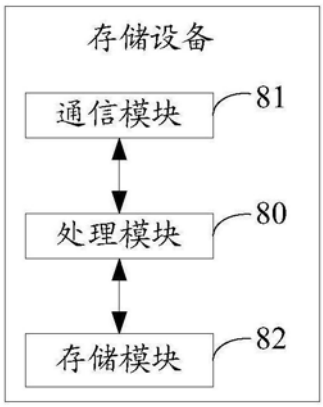

存储设备
通信模块
81
处理模块
80
存储模块
82

图8

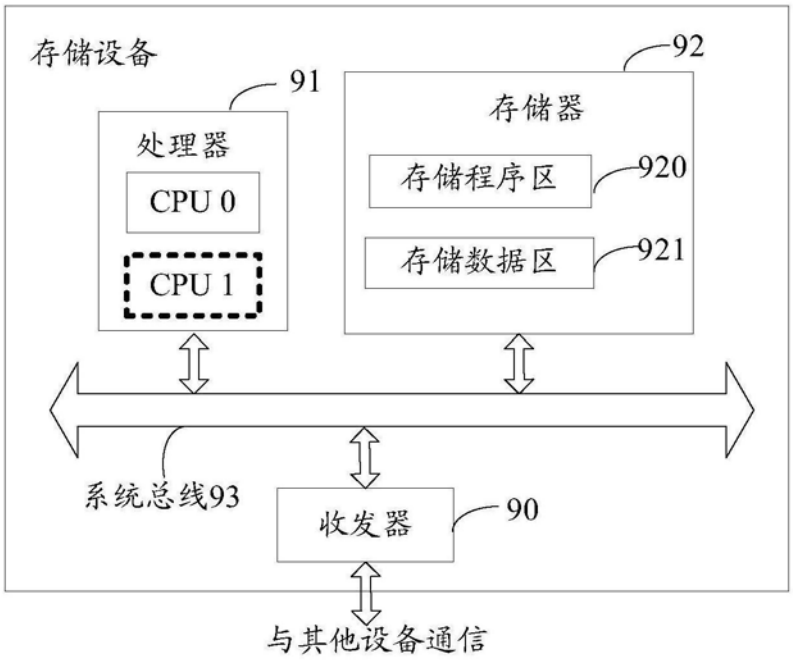

存储设备
91
处理器
CPU 0
CPU 1
92
存储器
存储程序区
920
存储数据区
921
系统总线93
收发器
90
与其他设备通信

图9