



(12)发明专利

(10)授权公告号 CN 104050073 B

(45)授权公告日 2018.02.27

(21)申请号 201410095467.9

(22)申请日 2014.03.14

(65)同一申请的已公布的文献号

申请公布号 CN 104050073 A

(43)申请公布日 2014.09.17

(30)优先权数据

13/843,869 2013.03.15 US

(73)专利权人 希捷科技有限公司

地址 美国加利福尼亚州

(72)发明人 M·A·盖尔特纳 K·道

S·法伍尔赫伯

(74)专利代理机构 上海专利商标事务所有限公

司 31100

代理人 何焜

(51)Int.Cl.

G06F 11/32(2006.01)

(56)对比文件

JP 特开2002-23966 A, 2002.01.25, 说明书
第26-165段.

审查员 陈沙沙

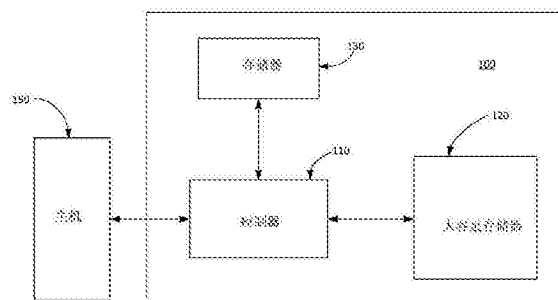
权利要求书2页 说明书5页 附图2页

(54)发明名称

暂态奇偶性/冗余

(57)摘要

本申请公开了暂态奇偶性/冗余。大容量存储使用额外的错误校正代码。额外的代码能够与关联数据分离地存储在存储介质(例如,易失性固态存储器)中。额外的代码可写到非易失性介质上。额外的代码可以是暂态的。额外的代码可以被高速缓存。只要存在,额外的代码可用于与其它的错误检测和校正代码同步地或者作为其它错误检测与校正代码的辅助对用户数据进行校正。



1. 一种用于数据存储的方法,包括:

在第一存储器中保持基于主机写的ECC,所述基于主机写的ECC是基于从主机写入的信息而生成的;以及

在与所述基于主机写的ECC相关联的数据被写入与所述第一存储器不同的第二存储器中后,响应于高速缓存策略而从所述第一存储器中逐出所述基于主机写的ECC。

2. 如权利要求1所述的方法,其中所述高速缓存策略是最不近期使用(LRU)、最不频繁使用(LFU)、先入先出(FIFO)和组相联中的一种。

3. 如权利要求1所述的方法,还包括使用所述基于主机写的ECC用于内部可靠性操作。

4. 如权利要求3所述的方法,其中所述内部可靠性操作是背景介质扫描、奇偶性生成或读后写验证中的一种。

5. 如权利要求1所述的方法,其中所述保持包括对所述基于主机写的ECC进行高速缓存。

6. 如权利要求1所述的方法,还包括暂时使用所述基于主机写的ECC。

7. 如权利要求6所述的方法,其中所述暂时使用是非单次使用。

8. 如权利要求6所述的方法,其中所述暂时使用是单次使用。

9. 如权利要求1所述的方法,还包括按与其它ECC结合使用、作为其它ECC的辅助或者作为其它ECC的备选使用中的至少一种方式来使用基于主机写的ECC。

10. 一种用于数据存储的方法,包括:

计算基于主机写的ECC,所述基于主机写的ECC是基于从主机写入的信息而生成的;

将所述基于主机写的ECC暂时保持在易失性存储器中;以及

在与所述基于主机写的ECC相关联的数据被写入与所述易失性存储器不同的存储器中后,响应于高速缓存策略而从所述易失性存储器中逐出所述基于主机写的ECC。

11. 如权利要求10所述的方法,其中所述基于主机写的ECC的暂时保持使用高速缓存策略。

12. 如权利要求11所述的方法,其中所述高速缓存策略是基于是否已生成基于主机写的ECC而改变的逐出策略。

13. 如权利要求10所述的方法,还包括将所述基于主机写的ECC和对于的元数据写入到非易失性存储器中。

14. 如权利要求10所述的方法,其中暂时保持所述基于主机写的ECC在一些地址范围内是优选进行的。

15. 如权利要求10所述的方法,还包括使用所述基于主机写的ECC用于恢复用户读的数据以及在内部操作期间恢复数据中的至少一项。

16. 如权利要求13所述的方法,暂时保持是非单次使用。

17. 大容量存储设备,包括:

第一存储器;

第二存储器;以及

控制器,其与所述第一存储器和所述第二存储器耦合,并且配置为将基于主机写的ECC暂时保持在至少所述第一存储器中,其中所述基于主机写的ECC是基于从主机写入的信息而生成的;并且所述控制器被进一步配置为在与所述基于主机写的ECC相关联的数据被写

入与所述第一存储器不同的第二存储器中后,响应于高速缓存策略而从所述第一存储器中逐出所述基于主机写的ECC。

18.如权利要求17所述的大容量存储设备,其中所述第一存储器易失性的,并且所述第二存储器是非易失性的。

19.如权利要求17所述的大容量存储设备,其中所述控制器使用高速缓存策略来暂时保持所述基于主机写的ECC。

20.如权利要求17所述的大容量存储设备,其中所述基于主机写的ECC是在关联数据被进行释放、删除和逐出中的一项之前被进行释放、删除和逐出中的一项。

暂态奇偶性/冗余

[0001] 发明概述

[0002] 大容量存储使用额外的错误校正代码。额外的代码能够与关联数据分离地存储在存储介质(例如,易失性固态存储器)中。额外的代码可写入非易失性介质中,可以是暂态的,或者可以进行高速缓存。只要存在,额外的代码就可用于与其它错误校正代码一起或作为其它错误校正代码的辅助对用户数据进行校正。

附图说明

[0003] 图1示出了使用额外错误校正代码的设备;

[0004] 图2示出了使用额外错误校正代码的另一设备;

[0005] 图3示出了使用额外错误校正代码的再一设备;

[0006] 图4示出了能够使用额外错误校正代码的数据;

[0007] 图5示出了使用额外错误校正代码的方法的流程图;以及

[0008] 图6示出了使用额外错误校正代码的大容量存储的层级关系。

[0009] 发明详述

[0010] 大容量存储持久地保存信息并且具有足够大以用于使用其的用户或系统的需求的容量。大容量存储包括如磁带库、RAID和JBOD的系统。大容量存储还包括如硬盘驱动器、磁盘驱动器、光盘驱动器、磁光驱动器和固态驱动器的设备。大容量存储还包括使用大容量存储系统或设备的基础结构。这些基础结构可以是直接附接的存储、网络附接的存储、存储域网、数据存储机架、分布式文件系统和云。

[0011] 大容量存储保持所保存信息的完整性的能力是所期望的特性。错误校正代码(ECC)的使用提供了保持信息完整性的方式。ECC随后能够与正在保持其完整性的信息一起被保存。这种保存ECC的方式具有成本。一方面成本是ECC需要存储空间。或者是用于保存信息的大容量存储内的空间进行重新分配以用于保存ECC,或者是提供额外的外部大容量存储来保存ECC。对于前者情况,大容量存储容量减小,而对于后者情况,预留空间增加了开销和复杂度。如果期望改进的或基于主机写的ECC,则大容量存储的容量将进一步减小,或者将提供额外的预留空间。任一种情况都加剧了所述成本。

[0012] 大容量存储能够通过保存信息的副本来保持所保存信息的完整性。所以,大容量存储不必使用ECC信息,这将消除关联的硬件/固件。然而,大容量存储的容量将必须大概加倍以保存既定量的信息或者如果不使用预留空间则将容量大概减半。而且,由于试图判定哪个副本是正确的而增加了复杂度。再有,使用该方案来保持所保存信息的完整性具有成本。

[0013] 另外,一些类型的ECC是在一组数据上计算出的。如果该组数据不完全写入,则存在与计算冗余相关的大的性能开销,因为需要从存储介质中读取全部的未写入数据来计算新的ECC。

[0014] 大容量存储能够通过使用基于主机写的ECC但是将其保存在与所保存信息所在的存储器或存储器的部分分离的存储器中来提高所保存信息的完整性。该基于主机写的ECC

覆盖了在单次主机传递中由数据所确定的存储器的可变范围。换言之,基于主机写的ECC未覆盖大容量存储器的单个单元(例如,扇区)或确定性的扇区组(例如,RAID奇偶性)。参考图1,示出了大容量存储设备100,其包括控制器110、大容量存储器120和存储器130。大容量存储器120可以包括至少一个磁盘、光盘、磁光盘、磁带、固态存储器或全息存储器、或其中至少两种的组合。存储器130可以包括易失性的或非易失性的存储器,诸如DRAM、SRAM、闪存或电池备用易失性存储器。控制器110可以使用存储器130作为用于待写入大容量存储器120的信息的高速缓存。

[0015] 主机190与大容量存储设备100耦合。主机190通过使用如SCSI、SAS、各种SATA(例如,mSATA、eSATA)、光纤通道等数据传输协议来与大容量存储设备100交互。主机190将信息写入大容量存储设备100,在大容量存储设备100中,信息由控制器110接收且被缓冲或高速缓存在存储器130中。然后,大容量存储设备100能够将信息保存到大容量存储器120中。当来自主机的信息正在存储器130之间传递时,大容量存储设备100能够生成基于主机写的ECC。基于主机写的ECC还能够在信息处于存储器130中的任意时间生成,例如,在从存储器130逐出之前。基于主机写的ECC包括任何错误校正代码,诸如Reed-Solomon、Hamming、BCH、Convolution、Turbo、简单奇偶性和LDPC。基于主机写的ECC能够包括诸如奇偶性、校验和和循环冗余校验等任何错误检测代码。

[0016] 基于主机写的ECC驻留在存储器130中。在被删除、释放或逐出之前,该基于主机写的ECC能够以各种方式使用。基于主机写的ECC能够用于确保用于用户或系统读操作的保存信息的完整性。基于主机写的ECC还能够用于确保在大容量存储设备100的内部可靠性操作期间所保持的信息的完整性,诸如背景介质扫描、带上的奇偶性生成或LBA的任何其它确定性的组,以及读后写验证。基于主机写的ECC

[0017] 基于多种标准中的一种标准的出现,基于主机写的ECC能够保持在存储器130中,从存储器130中删除、释放或逐出。如果通过高速缓存逐出策略来控制存储器130,则用于逐出基于主机写的ECC的标准能够响应于最不近期使用(LRU)、最不频繁使用(LFU)、FIFO、组相联等。此外,当基于主机写的ECC生成时,高速缓存策略可以改变以考虑到基于主机写的ECC的保持和释放。基于用于一些地址范围的优先权,诸如频繁写的地址或易出错地址,基于主机写的ECC能够保持在存储器130中,从存储器130中删除、释放或逐出。而且,基于主机写的ECC的未决逐出能够用于优先内部可靠性操作。此外,在基于主机写的ECC用于内部可靠性操作之后,基于主机写的ECC能够被删除、释放或逐出。并且,从大容量存储器120中逐出基于主机写的ECC能够在确保其如上所述使用之前被允许或者不被允许直到其如上所述使用。

[0018] 基于主机写的ECC能够以多种方式写入非易失性存储器。在图1中,基于主机写的ECC能够写入大容量存储器120,而不是存储器130中。例如,大容量存储器120可以是具有保持有基于主机写的ECC的介质高速缓存的磁盘。可选地,基于主机写的ECC能够在被写入存储器130中之后被写入大容量存储器120中。在该情况下,基于主机写的ECC到大容量存储器120的写能够在将其从存储器130逐出之前发生,或者当在大容量存储设备100中发生掉电时发生。对于掉电的情况,存储器130能够通过例如主轴电动机的备用EMF来供电,使得基于主机写的ECC能够被写入大容量存储器120中。

[0019] 利用如上所述用于存储器130的相同的标准,能够将额外的ECC保持在大容量存储

器120中,或者从大容量存储器120中删除、释放或逐出。在大容量存储器120中的基于主机写的ECC还能够如上所述用于存储器130。另外,从大容量存储器120中逐出基于主机写的ECC能够在确保其如上所述使用之前被允许或者不被允许直到其如上所述使用。

[0020] 基于主机写的ECC还能够具有关联的元数据。元数据是指描述基于主机写的ECC的数据。元数据可以包括基于主机写的ECC所覆盖的LBA范围、到基于主机写的ECC的位置的指针、到下一个最陈旧的基于主机写的ECC的链接、到下一个最年轻的基于主机写的ECC的链接、到覆盖下一个较高的LBA范围的基于主机写的ECC的链接以及到覆盖下一个较低的LBA范围的基于主机写的ECC的链接。对于每个基于主机写的ECC,元数据可以跟踪对应的LBA范围。例如,假设逻辑块地址100-109从主机或在存储器之间传递,并且生成基于主机写的ECC。元数据记录,对于那些逻辑块地址存在基于主机写的ECC,并且记录该基于主机写的ECC的位置(例如,DRAM地址)。还可能在不同类型的或配置的基于主机写的ECC。如果是这样,则也通过元数据进行描述。该元数据还能够存储在诸如大容量存储器120的非易失性存储器中,以使基于主机写的ECC在预期的和非预期的电力循环中被保留。

[0021] 存储器130还能够通过跟踪与先前存在的基于主机写的ECC重叠的基于主机写的ECC来保持有效的一组元数据。为了继续上述实施例,如果对于逻辑块地址100-109接收到另外的写(或某其它重叠方案),则先前存在的基于主机写的ECC现在无效,因为其覆盖了覆写数据。因此,应当适当地管理元数据以表明用于覆写数据的基于主机写的ECC是无效的且应当更新以描述新的基于主机写的ECC(如果有的话)。

[0022] 图2示出了使用基于主机写的ECC的大容量存储设备200。控制器210、大容量存储器220、易失性存储器235以及非易失性存储器240包括在大容量存储设备200中。大容量存储器220可以包括至少一个磁盘、光盘、磁光盘、磁带、固态存储器或全息存储器,或者其中至少两种的组合。存储器235可以包括诸如DRAM和SRAM的易失性存储器。非易失性存储器240可以包括诸如FLASH、STRAM、ReRAM、MRAM等非易失性固态存储器。主机290与大容量存储设备200耦合。

[0023] 大容量存储设备200能够将基于主机写的ECC保存在图示的存储器中的一种、两种或所有这三种中。例如,能够使用分层存储,其中首先将基于主机写的ECC保存在易失性存储器235中,然后保存到非易失性存储器240中,然后保存到大容量存储器220中。基于主机写的ECC在这些存储器之中的移动能够基于上述逐出标准。用于这些存储器的其它排列是可能的。基于上文针对图1中的大容量存储设备100所描述的任意方式,基于主机写的ECC能够写入到易失性存储器235、然后写入到非易失性存储器240中。基于主机写的ECC的保持、释放、删除、逐出和使用可与上文针对大容量存储设备100所述的相同。

[0024] 图3示出了与大容量存储设备300耦合的主机390,主机包括控制器310和存储器/大容量存储器325。存储器/大容量存储器325可以是同质存储介质,诸如快擦写存储器或磁盘。这种类型的存储介质既能够用作用于存储基于主机写的ECC的存储器,又能够用作用于保存关联信息的大容量存储器。如果使用快擦写存储器,则存储器/大容量存储器325的存储器部分可以是SLC,并且大容量存储器部分可以是MLC。另一选项是MLC设备的LSB页面。LSB页面典型地比MSB页面编程更快。如果使用磁盘,则存储器部分是介质高速缓存,并且大容量存储器部分可以是磁盘的其余表面。在任一情况下,存储器部分展示出比大容量存储器部分更快的读或写速度(或者两者)。基于主机写的ECC的维护、释放、删除、逐出和使用可

与上文针对大容量存储设备100所描述的相同。

[0025] 将参考图4来说明如此描述的基于主机写的ECC的应用。在图4中是DATA块410-470以及校正代码(ECC)块480。DATA块410-470中的每一个可以包括其自身的错误检测或校正信息(或两者)。ECC块480用于全部的DATA块410-470。块410-480可以在磁盘的磁道上、快擦写存储器的块上或RAID系统的带上是毗邻的,其中每个块410-480存储在不同的大容量存储设备上。

[0026] 与例如图1中的大容量存储设备100耦合的主机能够请求写信息到DATA块430-440。关联的基于主机写的ECC对于DATA块430-440生成,如上所述,并且保存在存储器中,诸如存储器130。EDC块480现在是无效的,因为其保护的数据的部分已经被重新写入。大容量存储设备100能够执行如背景介质扫描的内部可靠性操作。而且,ECC块480现在是无效的,因为其要保护的数据的部分已经重新写入。对于块430-440生成的ECC能够用于更新EDCECC块480。

[0027] 如果在DATA块430中发现错误,则大容量存储设备能够使用关联的基于主机写的ECC来与DATA块430中的错误检测或校正(或两者)信息或ECC块480(如果其已更新)结合地、作为DATA块430中的错误检测或校正(或两者)信息或ECC块480(如果其已更新)的辅助或备选来校正错误。

[0028] 图5示出了与所描述的基于主机写的ECC相关的方法500。方法500开始于步骤510,然后进行到步骤520,在该步骤中大容量存储保持基于主机写的ECC以校正信息错误。在步骤530中,对是否满足逐出、释放或删除标准做出判定。如果为否,则方法返回到步骤520。如果为是,则方法进行到步骤540,在该步骤中,基于主机写的ECC被逐出。方法500随后结束于步骤550。

[0029] 在上述任一种大容量存储设备中,基于主机写的ECC能够在关联信息被覆写、擦除或释放之前被逐出、释放或删除。然后,基于主机写的ECC相对于关联信息是暂态的。例如,大容量存储能够使用基于主机写的ECC来生成诸如EDC块480的奇偶性。然后,能够从存储器中逐出基于主机写的ECC。这是单次使用暂态性的实施例。作为另一实例,基于主机写的ECC能够如上所述使用,直到其被逐出、释放或删除。这是非单次使用暂态性的实施例。

[0030] 作为另外的解释,奇偶性/冗余计算能够在奇偶性带(即,RAID组)上进行且保存在与该奇偶性带耦合的地址位置上。上述大容量存储计算基于主机写的ECC,而无需获知任何奇偶性带的内容,或者甚至无需获知下面的奇偶性带的存在。基于主机写的ECC保持(有时是暂态地)在近期写入的数据上,而不一定在整个数据集(例如,奇偶性带)上。

[0031] 图6示出了如上所述使用基于主机写的ECC的大容量存储的层级关系。此处,基于主机写的ECC能够存储在大容量存储设备、系统或基础结构的存储器中。无论额外的ECC或所保存的信息驻留在何处,大容量存储设备、系统或基础结构能够取回信息和基于主机写的ECC并且如上所述使用其。例如,大容量存储系统620能够将基于主机写的ECC保持在其存储器中,然后将关联信息写入大容量存储设备610中。大容量存储系统620随后能够请求从大容量存储设备610读取关联信息。一旦大容量存储系统620接收到关联信息,其能够使用基于主机写的ECC与如上所述的关联信息。

[0032] 作为另一实施例,大容量存储系统620能够将基于主机写的ECC保持在其存储器中,然后,将关联信息写入大容量存储设备610中。大容量存储设备610稍后会从大容量存储

系统620请求基于主机写的ECC,并且如上所述对其进行使用。类似的交互能够与大容量存储基础结构630发生。

[0033] 上述方法和装置能够使用基于主机写的ECC来从可能未对近期写入数据造成可检测故障或任何其它错误的写入错误中恢复数据。控制器单独地或者与关联的固件一起来控制基于主机写的ECC在一个或多个存储器中的维护。

[0034] 上面的说明阐述了细节,但是不应视为限制权利要求的范围。对上述说明的各种修改、删除和添加都在权利要求的范围之内。

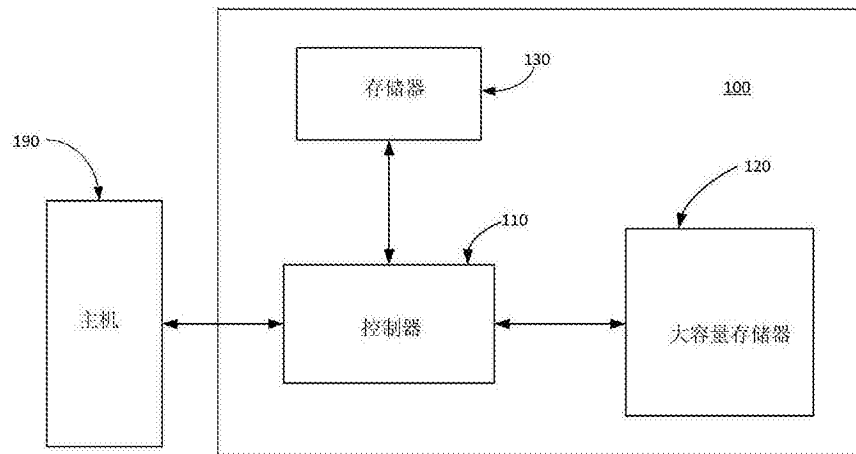


图1

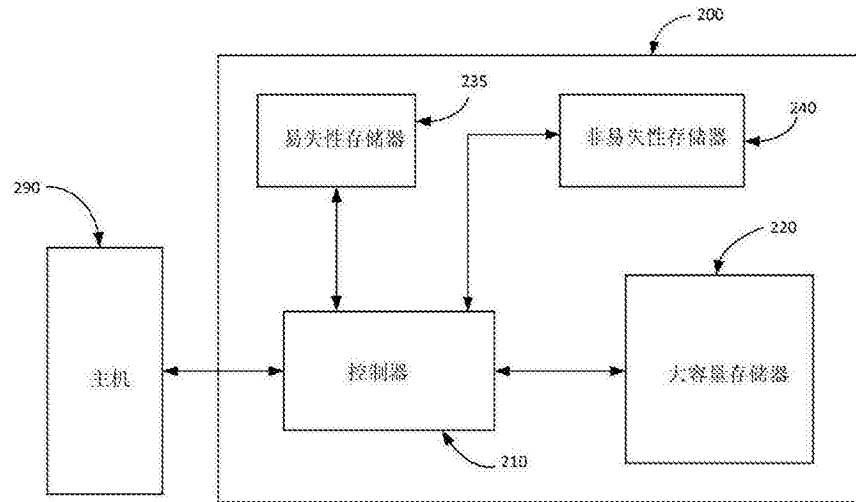


图2

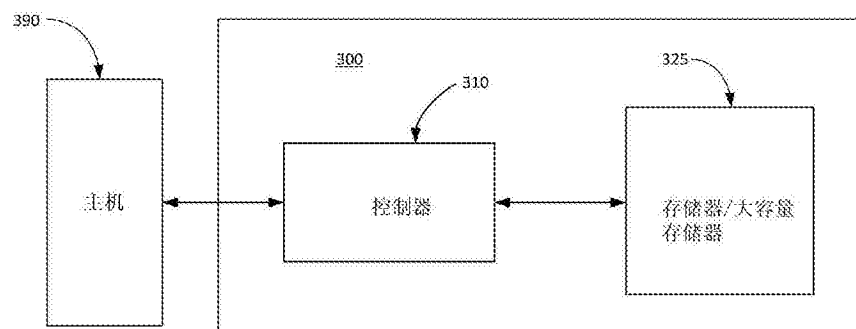


图3

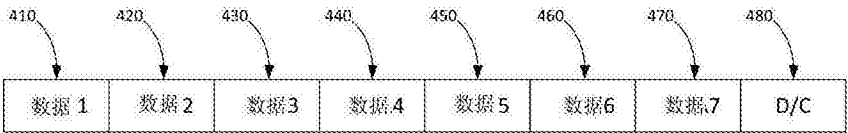


图4

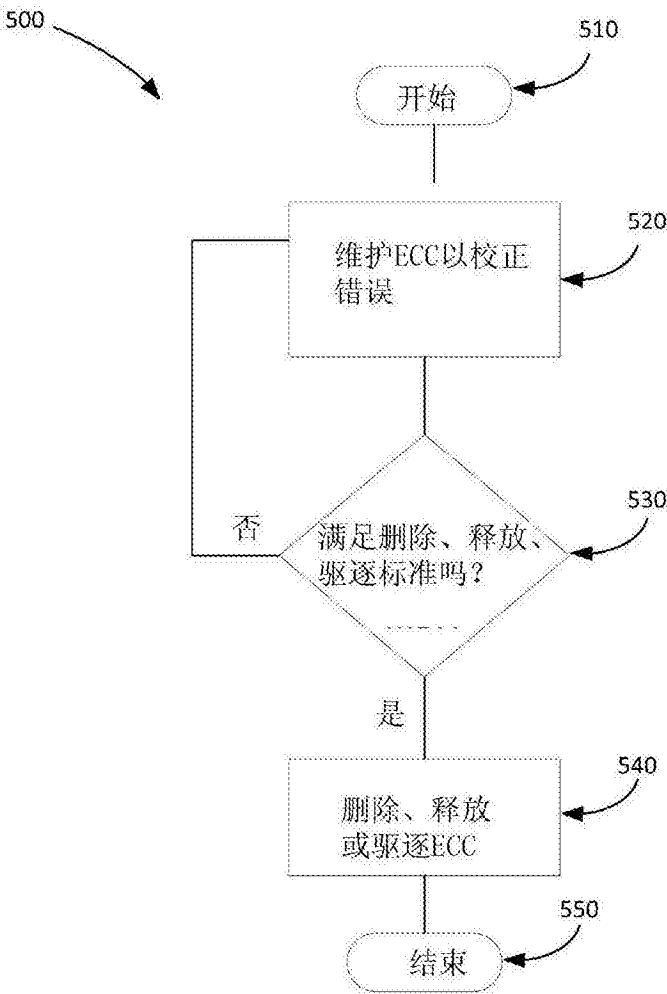


图5

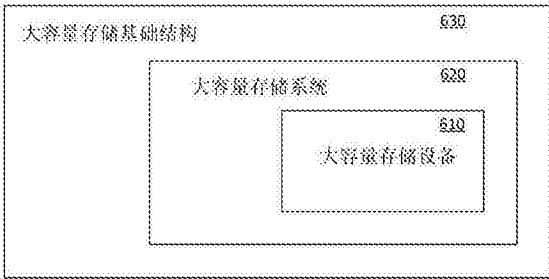


图6