



(12) 发明专利

(10) 授权公告号 CN 110612511 B

(45) 授权公告日 2023. 10. 03

(21) 申请号 201880028884.8

(22) 申请日 2018.03.09

(65) 同一申请的已公布的文献号
申请公布号 CN 110612511 A

(43) 申请公布日 2019.12.24

(30) 优先权数据
15/453,949 2017.03.09 US

(85) PCT国际申请进入国家阶段日
2019.10.31

(86) PCT国际申请的申请数据
PCT/US2018/021659 2018.03.09

(87) PCT国际申请的公布数据
W02018/165502 EN 2018.09.13

(73) 专利权人 NETAPP股份有限公司
地址 美国加利福尼亚州

(72) 发明人 R·德罗纳拉朱 K·D·斯特林
M·K·布哈塔查尔杰 M·古普塔

(74) 专利代理机构 北京市金杜律师事务所
11256
专利代理师 马明月

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

(56) 对比文件
CN 101495978 A, 2009.07.29
US 9471585 B1, 2016.10.18
US 2016313943 A1, 2016.10.27
US 2011107042 A1, 2011.05.05
CN 101663654 A, 2010.03.03
CN 105051675 A, 2015.11.11
CN 102939765 A, 2013.02.20

审查员 胡嫵

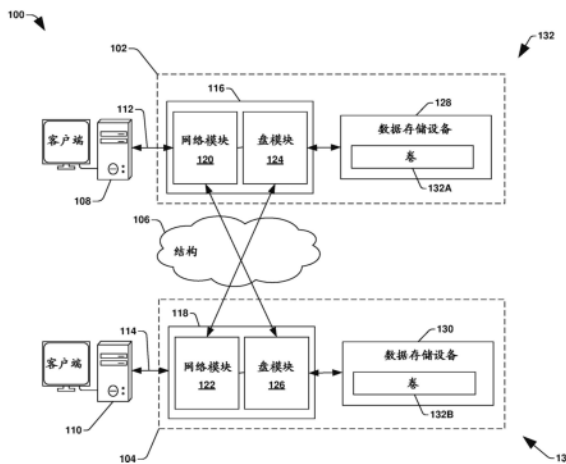
权利要求书3页 说明书14页 附图8页

(54) 发明名称

使用流选择性地分配区域中存储数据

(57) 摘要

提供了使用流将数据选择性地存储到分配区域中的技术。为存储设备定义了一组分配区域(例如,诸如虚拟块编号等块编号的范围)。向存储设备发送具有特定特性的数据(例如,用户数据、元数据、热数据、冷数据、随机访问数据、顺序访问数据等)以在相应的分配区域中进行选择性存储。例如,当文件系统接收到热数据的写入流时,可以将热数据指派给流。将使用流标识符对流进行标记,该流标识符用作存储设备的指示符,以使用为热数据定义的分配区域来处理流的数据。以这种方式,具有不同特性的数据将被存储/限制在存储设备的特定分配区域内,以减少碎片和写入放大。



1. 一种用于存储数据的方法,包括:

由文件系统执行与存储设备的协商以:指定利用第一流标识符而被标记的第一流的数据要被存储到所述存储设备的第一分配区域中,以及指定利用第二流标识符而被标记的第二流的数据要被存储到所述存储设备的第二分配区域中;

响应于接收第一数据以写入所述存储设备,基于与所述存储设备的分配区域的第一可用空闲空间相对应的第一排序标准、所述第一数据的第一特性、以及向所述分配区域指派特性的策略,将所述分配区域排序和排名成第一组经排序的分配区域;

基于所述第一分配区域具有高于第一阈值的第一排序排名,从所述第一组经排序的分配区域中选择第一分配区域;

响应于接收用以写入所述存储设备的第二数据,基于与所述分配区域的第二当前可用空闲空间相对应的第二排序标准、所述第二数据的第二特性、以及所述策略,将所述存储设备的所述分配区域重新排序和重新排名成第二组经排序的分配区域;

基于所述第二分配区域具有高于第二阈值的第二排序排名,从所述第二组经排序的分配区域中选择第二分配区域;以及

通过利用所述第一流标识符而被标记的所述第一流向所述存储设备发送所述第一数据,以及通过利用所述第二流标识符而被标记的所述第二流向所述存储设备发送所述第二数据。

2. 根据权利要求1所述的方法,其中策略指定流标识符以用于具有用户数据特性的数据和用于具有元数据特性的数据。

3. 根据权利要求1所述的方法,其中所述第一阈值或者所述第二阈值中的至少一个对应于最高排序排名。

4. 根据权利要求1所述的方法,其中所述第一分配区域和所述第二分配区域被定义在所述存储设备的单个存储介质内。

5. 根据权利要求1所述的方法,其中策略指定所述第一流标识符用于具有与所述存储设备相关联的第一聚合的第一存储聚合特性的数据,并且指定所述第二流标识符用于具有与所述存储设备相关联的第二聚合的第二存储聚合特性的数据。

6. 根据权利要求1所述的方法,其中所述文件系统将所述第一数据指派给所述第一流并且将所述第二数据指派给所述第二流。

7. 根据权利要求1所述的方法,其中所述存储设备由虚拟化层使用以提供对数据的访问。

8. 根据权利要求7所述的方法,其中在所述存储设备缺少逻辑地址到物理地址的映射时,所述虚拟化层提供对所述数据的访问。

9. 一种非暂态机器可读介质,包括用于执行方法的指令,所述指令在被机器执行时,使得所述机器执行以下操作:

由文件系统执行与存储设备的协商以:指定利用第一流标识符而被标记的第一流的数据要被存储到所述存储设备的第一分配区域内,以及指定利用第二流标识符而被标记的第二流的数据要被存储到所述存储设备的第二分配区域内;

响应于接收第一数据以写入所述存储设备,基于与所述存储设备的分配区域的第一可用空闲空间相对应的第一排序标准、所述第一数据的第一特性、以及向所述分配区域指派

特性的策略,将所述分配区域排序和排名成第一组经排序的分配区域;

基于所述第一分配区域具有高于第一阈值的第一排序排名,从所述第一组经排序的分配区域中选择第一分配区域;

响应于接收用以写入所述存储设备的第二数据,基于与所述分配区域的第二当前可用空闲空间相对应的第二排序标准、所述第二数据的第二特性、以及所述策略,将所述存储设备的所述分配区域重新排序和重新排名成第二组经排序的分配区域;

基于所述第二分配区域具有高于第二阈值的第二排序排名,从所述第二组经排序的分配区域中选择第二分配区域;以及

通过利用所述第一流标识符而被标记的所述第一流向所述存储设备发送所述第一数据,以及通过利用所述第二流标识符而被标记的所述第二流向所述存储设备发送所述第二数据。

10. 根据权利要求9所述的非暂态机器可读介质,其中所述指令使得所述机器:

将所述第一分配区域和所述第二分配区域定义为所述存储设备的擦除块单元的整数倍。

11. 根据权利要求9所述的非暂态机器可读介质,其中策略指示随机访问数据要被存储在所述第一分配区域内。

12. 根据权利要求9所述的非暂态机器可读介质,策略指示顺序访问数据要被存储在所述第二分配区域内。

13. 根据权利要求9所述的非暂态机器可读介质,其中所述第一阈值或者所述第二阈值中的至少一个对应于最高排序排名。

14. 根据权利要求9所述的非暂态机器可读介质,其中所述第一分配区域和所述第二分配区域被定义在所述存储设备的单个存储介质内。

15. 根据权利要求9所述的非暂态机器可读介质,其中所述指令使得所述机器:

基于确定所述第一分配区域具有小于阈值的空闲空间量,重新定义所述第一分配区域以增加所述第一分配区域的可用存储空间。

16. 一种用于数据存储的计算设备,包括:

存储器,包括机器可执行代码;以及

被耦合到所述存储器的处理器,所述处理器被配置为执行所述机器可执行代码以使得所述处理器执行:

由文件系统执行与存储设备的协商以:指定利用第一流标识符而被标记的第一流的数据要被存储到所述存储设备的第一分配区域中,以及指定利用第二流标识符而被标记的第二流的数据要被存储到所述存储设备的第二分配区域中;

响应于接收第一数据以写入所述存储设备,基于与所述存储设备的分配区域的第一可用空闲空间相对应的第一排序标准、所述第一数据的第一特性、以及向所述分配区域指派特性的策略,将所述分配区域排序和排名成第一组经排序的分配区域;

基于所述第一分配区域具有高于第一阈值的第一排序排名,从所述第一组经排序的分配区域中选择第一分配区域;

响应于接收用以写入所述存储设备的第二数据,基于与所述分配区域的第二当前可用空闲空间相对应的第二排序标准、所述第二数据的第二特性、以及所述策略,将所述存储设

备的所述分配区域重新排序和重新排名成第二组经排序的分配区域；

基于所述第二分配区域具有高于第二阈值的第二排序排名，从所述第二组经排序的分配区域中选择第二分配区域；以及

通过利用所述第一流标识符而被标记的所述第一流向所述存储设备发送所述第一数据，以及通过利用所述第二流标识符而被标记的所述第二流向所述存储设备发送所述第二数据。

17. 根据权利要求16所述的计算设备，其中所述存储设备是具有第一分区和第二分区的分区存储设备。

18. 根据权利要求16所述的计算设备，其中所述第一分配区域和所述第二分配区域被定义在所述存储设备的单个存储介质内。

19. 根据权利要求16所述的计算设备，其中基于所述第二数据被存储在一组不连续的块内，所述第二数据被标识为随机访问数据。

20. 根据权利要求16所述的计算设备，其中基于所述第一数据被存储在一组连续的块内，所述第一数据被标识为顺序访问数据。

使用流选择性地向分配区域中存储数据

背景技术

[0001] 存储控制器的很多文件系统可以通过与各种协议相关联的多个通道来接收数据。文件系统可以将该数据存储存储在存储设备内或跨多个存储设备存储。这样的数据可以具有各种特性,例如诸如是用户数据(例如,用户数据库文件)或元数据(例如,卷的卷大小、存储控制器的网络地址、复制策略、和/或由文件系统和/或存储控制器使用的其他数据)。在另一示例中,特性可以对应于热数据(例如,正在以高于阈值频率的频率被访问的数据,诸如被文件系统频繁修改的元数据)或冷数据(例如,正在以低于阈值频率的频率被访问的用户数据)。在另一示例中,特性可以对应于顺序地被访问的数据(例如,存储在连续块内的数据)或随机地被访问的数据(例如,存储在不连续的块内的数据)。

[0002] 虚拟化层可以被用作间接层,该间接层将来自多个存储设备的物理存储组合在一起为在客户端和应用看起来是单个存储对象(例如,卷或逻辑单元号(LUN),其可以跨越多个物理存储设备)。虚拟化层抽象化存储的物理布局,并且因此在被映射到底层物理地址空间的逻辑地址空间中操作。利用虚拟化存储,存储设备可以承担将数据物理地存储在存储设备的物理块内的由存储设备选择的位置中的作用。然而,存储设备可能将任何类型的数据(诸如热数据、冷数据、用户数据和元数据)存储在一起,而不进行任何逻辑/物理分隔。不幸的是,具有不同特性的数据可能具有不同的访问和覆写模式,并且因此,当这样的数据被存储在一起时会产生碎片。

[0003] 在一个示例中,固态驱动器(SSD)可能不具有覆写先前写入的块的能力,而该SSD只能向空的目的地单元写入。因此,如果目的地单元具有要被重用的数据,则该数据可以被移动到不同的空单元并且目的地单元必须被重新编程(例如,擦除)以便可以将新数据写入目的地单元。这导致碎片增加,而当将不同特性的数据存储在一起时,碎片增加变得更加明显。存储设备可以保留空间以提供可以前摄地释放单元的后台垃圾回收(background garbage collection)。但是,诸如存储设备的大约28%或任何其他百分比的存储的大量存储空间可能被保留。由于过度配置,这会导致存储资源的低效使用和成本增加。而且,写入放大在随后的覆写方面变得成问题,这可能导致性能下降和存储设备的磨损。

附图说明

[0004] 图1是示出根据本文中阐述的一个或多个规定的示例集群网络的组件框图。

[0005] 图2是示出根据本文中阐述的一个或多个规定的示例数据存储系统的组件框图。

[0006] 图3是示出使用流选择性地向分配区域中存储数据的示例性方法的流程图。

[0007] 图4是示出用于使用流选择性地向分配区域中存储数据的示例性计算设备的组件框图,其中分配区域被定义并且策略被指派给分配区域。

[0008] 图5是示出用于使用流选择性地向分配区域中存储数据的示例性计算设备的组件框图,其中分配区域跨多个存储设备被定义。

[0009] 图6是示出使用流选择性地向分配区域中存储数据的示例性方法的流程图。

[0010] 图7是示出用于使用流选择性地向分配区域中存储数据的示例性计算设备的组件

框图。

[0011] 图8是根据本文中阐述的一个或多个规定的计算机可读介质的一个示例。

具体实施方式

[0012] 现在参考附图描述所要求保护的主体的一些示例,其中相同的附图标记通常用于始终表示相同的元件。在以下描述中,出于解释的目的,阐述了很多具体细节以便提供对所要求保护的主体理解。然而,清楚的是,可以在没有这些具体细节的情况下实践所要求保护的主体。本具体实施方式中的任何内容均不被承认为是现有技术。

[0013] 本文中提供了用于使用流选择性地分配区域中存储数据的一种或多种技术和/或计算设备。

[0014] 存储设备可以由虚拟化层使用以向客户端提供虚拟化存储(例如,虚拟化层可以隐藏物理存储的底层细节,并且可以将多个物理设备的物理存储组合为暴露于客户端和应用的单个存储对象)。如果存储设备不具有逻辑地址空间到物理地址空间的明确定义的映射,则存储设备将仅将任何类型的数据存储在一起。将不同类型的数据(例如,具有不同访问频率的数据、具有不同覆写模式和频率的数据、不同聚合的数据、随机访问数据、顺序访问数据、热数据、冷数据、用户数据、元数据等)存储在一起久而久之能够导致存储设备的碎片化。写入放大也将影响后续的覆写。而且,诸如固态驱动器等某些类型的存储设备过度提供存储(例如,保留一定比例的空闲存储),该存储被垃圾回收功能使用以前摄地释放固态驱动器的单元。这样的过度配置浪费了原本可以被用来存储用户数据和/或元数据的存储空间。因此,需要减少碎片化,减少写入放大,并且提高存储空间的存储效率。

[0015] 因此,如本文中提供的,基于写入流的数据的特性将这样的数据指派(assign)给不同的流。特别地,基于指定具有不同特性的数据将要被存储在存储设备的不同分配区域中(例如,在存储设备的不同物理地址范围或存储设备的虚拟块编号内)的一个或多个策略,可以将频繁访问的数据指派给第一流,可以将不频繁访问的数据指派给第二流,可以将随机访问数据指派给第三流,可以将顺序访问数据指派给第四流,等等。

[0016] 策略和数据到流的分配可以由文件系统实现,使得即使在存储设备的存储被虚拟化和/或存储设备没有维持逻辑地址空间到物理地址空间的明确定义的映射,并且因此所有数据仅被存储在一起或没有任何区别地被存储的情况下,数据能够存储在存储设备内的分离的位置。可以使用由策略指派给相应分配区域的特定流标识符来标记每个流。例如,该策略可以指定频繁访问的数据是要被存储在分配区域(C)中,并且因此,频繁访问的数据的第一流用如下流标识符来标记:该流标识符被用作给存储设备的、第一流的数据将要使用分配区域(C)来处理(例如,被存储在其中)的指示符。将具有相似特性的数据一起存储在同一个分配区域中并且将具有不相似特性的数据存储在不同的分配区域中能为整个存储设备减少碎片化和写入放大(例如,频繁地覆写的数据可以被包含在单个分配区域中,而不是跨整个存储设备而散布,使得来自覆写的碎片将不会影响整个存储设备,否则,会在整个存储设备上产生碎片,特别是对于随处写入(write anywhere)文件系统,其在任何写入操作中将数据写入新位置)。因为可以不需要后台垃圾回收过程,或者可以使用存储设备的小得多的保留区域用于垃圾回收,这还提高了存储效率。

[0017] 为了提供使用流选择性地分配区域中存储数据,图1示出了集群网络环境100或

网络存储环境的一个实施例。然而,可以理解,本文中描述的技术等可以在集群网络环境100、非集群网络环境和/或各种其他计算环境(诸如台式计算环境)内实现。也就是说,包括所附权利要求的范围在内的本公开内容并不意味着被限于本文中提供的示例。应当理解,如果在后的附图中示出了关于在先的附图先前讨论过的相同或相似的组件、元件、特征、项目、模块等,则在描述后续附图时可以省略(例如,为了简化和易于理解)对于这些内容的类似的(例如,冗余的)讨论。

[0018] 图1是示出可以实现本文中描述的技术和/或系统的至少一些实施例的集群网络环境100的框图。集群网络环境100包括通过集群结构106耦合的数据存储系统102和104,集群结构106诸如实施为私有无限带宽(Infiniband)、光纤通道(FC)或以太网网络的计算网络,该计算网络促进数据存储系统102和104(以及其中的一个或多个模块、组件等,例如,诸如节点116和118)之间的通信。应当理解,虽然图1中示出了两个数据存储系统102和104以及两个节点116和118,但是可以预期这样的组件的任何合适数目。在一个示例中,节点116、118包括存储控制器(例如,节点116可以包括主存储控制器或本地存储控制器,并且节点118可以包括辅助存储控制器或远程存储控制器),这些存储控制器向诸如主机设备108、110等客户端设备提供对存储在数据存储设备128、130内的数据的访问。类似地,除非本文中另外特别地提供,否则对于本文中引用和/或在附图中示出的其他模块、元件、特征、项目等也是同样如此。也就是说,本文中公开的组件、模块、元件、特征、项目等的特定数目不意在以限制方式来解释。

[0019] 可以进一步理解,集群网络不限于任何特定地理区域,并且可以是本地地和/或远程地群集的。因此,在一个实施例中,集群网络可以分布在位于多个地理位置的多个存储系统和/或节点上;而在另一实施例中,集群网络可以包括驻留在相同地理位置(例如,在数据存储设备的单个现场机架)中的数据存储系统(例如,102、104)。

[0020] 在所示例中,一个或多个主机设备108、110可以包括例如:客户端设备、个人计算机(PC)、用于存储的计算设备(例如,存储服务器)、以及其他计算机或外围设备(例如,打印机),该一个或多个主机设备通过存储网络连接112、114被耦合到相应的数据存储系统102、104。网络连接可以包括例如利用网络附属存储(NAS)协议(诸如用以交换数据分组的通用因特网文件系统(CIFS)协议或网络文件系统(NFS)协议)、存储区域网络(SAN)协议(诸如小型计算机系统接口(SCSI)或光纤通道协议(FCP))、对象协议(诸如简单存储服务S3等)的局域网(LAN)或广域网(WAN)。说明性地,主机设备108、110可以是运行应用的通用计算机,并且可以使用用于信息交换的客户端/服务器模型与数据存储系统102、104交互。也就是说,主机设备可以从数据存储系统请求数据(例如,由网络存储控制器管理的存储设备上的数据,该网络存储控制器被配置为处理由主机设备针对存储设备发出的I/O命令),并且数据存储系统可以经由一个或多个存储网络连接112、114将该请求的结果返回给主机设备。

[0021] 例如,集群数据存储系统102、104上的节点116、118可以包括被互连为集群的网络或主机节点,以向诸如具有远程位置的企业、云存储(例如,存储端点可以存储在数据云内)等提供数据。例如,集群网络环境100中的这样的节点可以是作为连接点、再指派点或通信端点而被附接到网络的设备。节点可以能够通过网络通信信道发送、接收和/或转发信息,并且可以包括满足这些标准中的任何或所有标准的任何设备。节点的一个示例可以是附接

到网络的数据存储和管理服务器,其中服务器可以包括通用计算机或特别地配置成操作为数据存储和管理系统中的服务器的计算设备。

[0022] 在一个示例中,诸如节点116、118等第一节点集群(例如,被配置为提供对第一存储聚合的访问的第一组存储控制器,第一存储聚合包括一个或多个存储设备的第一逻辑分组)可以位于第一存储站点上。未示出的第二节点集群可以位于第二存储站点(例如,被配置为提供对第二存储聚合的访问的第二组存储控制器,第二存储聚合包括一个或多个存储设备的第二逻辑分组)上。第一节点集群和第二节点集群可以根据灾难恢复配置来配置,其中在包括灾难节点集群的灾难存储站点处发生灾难的情况下,幸存的节点集群提供对灾难节点集群的存储设备的切换访问(例如,在第二存储站点处发生灾难的情况下,第一节点集群向客户端设备提供对第二存储聚合的存储设备的切换数据访问)。

[0023] 如集群网络环境100中所示,节点116、118可以包括协调以为集群提供分布式存储架构的各种功能组件。例如,节点可以包括网络模块120、122和盘模块124、126。网络模块120、122可以被配置为允许节点116、118(例如,网络存储控制器)通过存储网络连接112、114与主机设备108、110连接,例如,以允许主机设备108、110访问存储在分布式存储系统中的数据。进一步地,网络模块120、122可以通过集群结构106提供与一个或多个其他组件的连接。例如,在图1中,节点116的网络模块120可以通过经由节点118的盘模块126发送请求来访问第二数据存储设备。

[0024] 盘模块124、126可以被配置为将一个或多个数据存储设备128、130(诸如盘或盘阵列、闪存或某种其他形式的数据存储)连接到节点116、118。节点116、118可以由集群结构106互连,例如,以允许集群中的相应的节点访问连接到集群中的不同节点的数据存储设备128、130上的数据。通常,盘模块124、126根据诸如SCSI或FCP等SAN协议与数据存储设备128、130通信。因此,从节点116、118上的操作系统来看,数据存储设备128、130可以表现为本地地附接到该操作系统。以这种方式,不同的节点116、118等可以通过操作系统来访问数据块,而无需通过特意地请求摘要文件(abstract file)。

[0025] 应当理解,虽然集群网络环境100示出了相同数目的网络和盘模块,但是其他实施例可以包括不同数目的这些模块。例如,集群中可以存在多个互连的网络和盘模块,而网络模块和盘模块之间不具有一一对应的关系。也就是说,不同的节点可以具有不同数目的网络模块和盘模块,并且同一节点可以具有与盘模块数目不同的网络模块。

[0026] 此外,主机设备108、110可以通过存储网络连接112、114与集群中的节点116、118联网。作为示例,被联网到集群的各个主机设备108、110可以请求集群中的节点116、118的服务(例如,以数据分组的形式的信息交换),并且节点116、118可以将所请求的服务的结果返回到主机设备108、110。在一个实施例中,主机设备108、110可以与驻留在数据存储系统102、104中的节点116、118(例如,网络主机)中的网络模块120、122交换信息。

[0027] 在一个实施例中,数据存储设备128、130包括卷132,卷132例如是将盘驱动器或盘阵列或其他存储(例如,闪存)上的信息的存储作为数据的文件系统的一个实现。在一个示例中,盘阵列可以包括全传统硬盘驱动器、全闪存驱动器、或传统硬盘驱动器和闪存驱动器的组合。例如,卷可以跨越盘的部分、盘的集合或多个盘的部分,并且通常地定义存储系统中的盘空间上的文件存储的整体逻辑布置。在一个实施例中,卷可以包括作为驻留在该卷内的分层目录结构中的一个或多个文件而存储的数据。

[0028] 卷通常地以可以与特定存储系统相关联的格式来配置,并且相应的卷格式通常包括向卷提供功能的特征,诸如向卷提供用以形成集群的能力。例如,在第一存储系统可以对于其卷利用第一格式时,第二存储系统可以对于其卷利用第二格式。

[0029] 在集群网络环境100中,主机设备108、110可以利用数据存储系统102、104来存储和取回卷132中的数据。在该实施例中,例如,主机设备108可以向数据存储系统102内的节点116中的网络模块120发送数据分组。节点116可以使用盘模块124向数据存储设备128转发数据,其中数据存储设备128包括卷132A。以这种方式,在该示例中,主机设备可以使用由存储网络连接112连接的数据存储系统102来访问卷132A,以存储和/或取回数据。此外,在该实施例中,主机设备110可以与数据存储系统104内的节点118中的网络模块122(例如,其可以远离数据存储系统102)交换数据。节点118可以使用盘模块126向数据存储设备130转发数据,从而访问与数据存储设备130相关联的卷132B。

[0030] 可以理解,可以在集群网络环境100内实现使用流选择性地向分配区域中存储数据。在一个示例中,可以在数据存储设备128和/或数据存储设备130内定义分配区域。可以通过流选择性地向数据存储设备128和/或数据存储设备130发送数据。可以用与要被处理的流所来自的分配区域相对应的流标识符来标记这样的流。可以理解,使用流选择性地向分配区域中存储数据可以在任何类型的计算环境中/或在任何类型的计算环境之间实现,并且在物理设备(例如,节点116、节点118、台式计算机、平板电脑、膝上型计算机、可穿戴设备、移动设备、存储设备、服务器等)和/或云计算环境(例如,相对集群网络环境100远程)之间可以是可转移的。

[0031] 图2是数据存储系统200(例如,图1中的102、104)的说明性示例,其提供了可以实现本文中描述的技术和/或系统中的一个或多个的组件的一个实施例的进一步细节。数据存储系统200包括节点202(例如,图1中的节点116、118)和数据存储设备234(例如,图1中的数据存储设备128、130)。节点202可以是例如通用计算机或者特别地配置成操作为存储服务器的某种其他计算设备。主机设备205(例如,图1中的108、110)可以例如通过网络216连接到节点202,以提供对存储在数据存储设备234上的文件和/或其他数据的访问。在一个示例中,节点202包括向客户端设备(诸如主机设备205)提供对存储在数据存储设备234内的数据的访问的存储控制器。

[0032] 数据存储设备234可以包括大容量存储设备,诸如盘阵列218、220、222的盘224、226、228。可以理解,本文中描述的技术和系统不受示例实施例的限制。例如,盘224、226、228可以包括任何类型的大容量存储设备,包括但不限于磁盘驱动器、闪存、以及适于存储信息(例如包括数据(D)和/或奇偶校验(P)信息)的任何其他类似介质。

[0033] 节点202包括一个或多个处理器204、存储器206、网络适配器210、集群访问适配器212和存储适配器214,其通过系统总线242而互连。数据存储系统200还包括安装在节点202的存储器206中的操作系统208,操作系统208可以例如实现独立(或廉价)磁盘冗余阵列(RAID)优化技术以优化阵列中的故障盘的数据的重建过程。

[0034] 操作系统208还可以管理针对数据存储系统的通信以及管理可以在集群网络中的诸如附接到集群结构215(例如,图1中的106)的其他数据存储系统之间的通信。因此,节点202(诸如网络存储控制器)可以响应于主机设备请求以根据这些主机设备请求来管理数据存储设备234(例如,或附加的集群设备)上的数据。操作系统208通常可以在数据存储系统

200上建立一个或多个文件系统,其中文件系统可以例如包括实现文件和目录的持久分层命名空间的软件代码和数据结构。作为一个示例,当将新的数据存储设备(未示出)添加到集群网络系统时,向操作系统208通知与新的数据存储设备相关联的新的文件要被存储在现有目录树中的何处。这通常被称为“安装”文件系统。

[0035] 在示例数据存储系统200中,存储器206可以包括由处理器204和适配器210、212、214可寻址用于存储相关的软件应用代码和数据结构的存储位置。例如,处理器204和适配器210、212、214可以包括被配置为执行软件代码并且操纵数据结构的处理元件和/或逻辑电路。操作系统208的部分通常驻留在存储器206中并且由处理元件执行,该操作系统通过尤其地调用存储操作以支持由存储系统实现的文件服务来在功能上组织存储系统。对于本领域技术人员来说很清楚的是,包括各种计算机可读介质的其他处理和存储机制可以用于存储和/或执行与本文中描述的技术有关的应用指令。例如,操作系统还可以利用一个或多个控制文件(未示出)来帮助虚拟机的提供。

[0036] 网络适配器210包括通过网络216将数据存储系统200连接到主机设备205所需要的机械、电气和信令电路,网络216可以尤其地包括点对点连接或诸如局域网的共享介质。主机设备205(例如,图1的108、110)可以是配置为执行应用的通用计算机。如上所描述,主机设备205可以根据信息递送的客户端/主机模型与数据存储系统200交互。

[0037] 存储适配器214与在节点202上执行的操作系统208协作以访问由主机设备205请求的信息(例如,访问由网络存储控制器管理的存储设备上的数据)。该信息可以存储在任何类型的可写介质的附接阵列上,可写介质诸如磁盘驱动器、闪存和/或适于存储信息的任何其他类似介质。在示例数据存储系统200中,信息可以存储在盘224、226、228上的数据块中。存储适配器214可以包括输入/输出(I/O)接口电路,其通过诸如存储区域网络(SAN)协议(例如,小型计算机系统接口(SCSI)、iSCSI、超SCSI(hyperSCSI)、光纤通道协议(FCP))的I/O互连布置耦合到盘。该信息由存储适配器214取回,并且如果需要的话,在通过系统总线242转发到网络适配器210(和/或在发送到集群中的另一节点的情况下的集群访问适配器212)之前,该信息由一个或多个处理器204(或存储适配器214本身)处理,其中该信息被格式化为数据分组,并且通过网络216返回到主机设备205(和/或通过集群结构215返回到附接到集群的另一节点)。

[0038] 在一个实施例中,盘阵列218、220、222上的信息存储可以实现为一个或多个存储卷230、232,这些存储卷包括定义盘空间的整体逻辑布置的盘集群224、226、228。包括一个或多个卷的盘224、226、228通常被组织为一或多个RAID组。作为一个示例,卷230包括盘阵列218和220的聚合,其包括盘集群224和226。

[0039] 在一个实施例中,为了便于访问盘224、226、228,操作系统208可以实现文件系统(例如,随处写入文件系统),该文件系统逻辑地将信息组织为盘上的目录和文件的分层结构。在该实施例中,相应的文件可以实现为被配置为存储信息的一组盘块,而目录可以实现为特定地格式化的文件,其中存储有关于其他文件和目录的信息。

[0040] 无论该数据存储系统200内的底层物理配置如何,数据都可以作为文件存储在物理和/或虚拟卷内,这些卷可以与诸如文件系统标识符(FSID)的相应的卷标识符相关联,这些标识符的长度在一个示例中可以是32位。

[0041] 物理卷对应于其地址、可寻址空间、位置等不改变的物理存储设备的至少部分,诸

如一个或多个数据存储设备234中(例如,独立(或廉价)磁盘冗余阵列(RAID系统))的至少一些数据存储设备。通常地,因为用于访问其的地址(范围)一般维持不变,物理卷的位置也不改变。

[0042] 相反地,虚拟卷存储在不同物理存储设备的不同部分的聚合上。虚拟卷可以是不同物理存储设备位置的不同可用部分的集合,诸如来自盘224、226和/或228中的每个盘的一些可用空间。可以理解,因为虚拟卷没有“绑定”到任何一个特定存储设备,可以说虚拟卷包括抽象或虚拟化层,这允许虚拟卷重新调整大小和/或在一些方面是灵活的。

[0043] 此外,虚拟卷可以包括一个或多个逻辑单元号(LUN)238、目录236、Q树(Qtree)235和文件240。尤其地,这些特征、但更特别地是LUN允许例如其内存储有数据的不同存储器位置被标识,并且被组合为数据存储单元。这样,LUN 238可以以构成的虚拟盘或驱动器为特点,在该虚拟盘或驱动器上虚拟卷内的数据被存储在聚合内。例如,LUN通常被称为虚拟驱动器,使得它们模仿来自通用计算机的硬盘驱动器,而LUN实际上包括存储在卷的各个部分中的数据块。

[0044] 在一个实施例中,一个或多个数据存储设备234可以具有一个或多个物理端口,其中每个物理端口可以被指派目标地址(例如,SCSI目标地址)。为了表示存储在数据存储设备上的相应的卷,数据存储设备上的目标地址可以被用于标识一个或多个LUN 238。因此,例如,当节点202通过存储适配器214连接到卷230、232时,创建节点202与该卷下的一个或多个LUN 238之间的连接。

[0045] 在一个实施例中,相应的目标地址可以标识多个LUN,使得目标地址可以表示多个卷。I/O接口可以例如实现为存储适配器214中的电路和/或软件或者驻留在存储器206中并且由处理器204执行的可执行代码,该I/O接口可以通过使用标识一个或多个LUN 238的一个或多个地址连接到卷230。

[0046] 可以理解,可以针对数据存储系统200实现使用流选择性地向分配区域中存储数据。在一个示例中,可以在一个或多个数据存储设备234内定义分配区域。可以通过流选择性地向一个或多个数据存储设备234发送数据。可以用与要被处理的流所来自的分配区域相对应的流标识符来标记这样的流。可以理解,使用流选择性地向分配区域中存储数据可以在任何类型的计算环境中和/或在任何类型的计算环境之间实现,并且在物理设备(例如,节点202、主机设备205、台式计算机、平板电脑、膝上型计算机、可穿戴设备、移动设备、存储设备、服务器等)和/或云计算环境(例如,对节点202和/或主机设备205是远程的)之间可以是可转移的。

[0047] 图3的示例性方法300示出了使用流选择性地向分配区域中存储数据的一个实施例。在302处,可以将存储设备的第一区域定义为第一分配区域。可以将存储设备的第二区域定义为第二分配区域。可以理解,可以为存储设备定义任何数目的分配区域和/或单个分配区域可以跨越任何数目和类型的存储设备(例如,跨越第一存储介质的第一部分、第二存储介质的第二部分等的分配区域;跨越RAID配置的第一盘、第二盘和奇偶校验盘的分配区域;等等)。该存储设备可以包括任何类型的存储设备,诸如固态设备、闪存设备、分区存储设备、缺少逻辑地址空间到物理地址空间的直接映射的存储设备、由诸如虚拟化存储设备的存储的虚拟化层的间接层使用的存储设备等。在一个示例中,分配区域可以被定义为诸如固态设备的存储设备的擦除块单元的整数倍(例如,数据以包括有多个单元的页面单元

被写入闪存,并且闪存只能以被称为块单元的较大单元被擦除,块单元包括多个页面单元)。

[0048] 在304处,可以(例如,由文件系统)促进与存储设备的协商以指定第一流标识符将被用作存储设备的第一指示符以指示用第一流标识符标记的流将要使用第一分配区域来处理(例如,用第一流标识符标记的流的数据将由存储设备存储在第一分配区域中,而不存储在其他分配区域中)。在306处,可以(例如,由文件系统)促进与存储设备的协商以指定第二流标识符将被用作存储设备的第二指示符以指示用第二流标识符标记的流将要使用第二分配区域来处理。以这种方式,存储设备将同意使用与流标识符相对应的分配区域来处理这样的流,该标识符由文件系统用于标记这样的流。

[0049] 在308处,可以维持指定具有某些特性的数据将要使用某些分配区域来处理的策略。例如,策略可以指定具有第一特性的数据将要使用第一分配区域来处理(例如,这样的数据将要从第一分配区域存储和读取),以及具有第二特性的数据将要使用第二分配区域来处理(例如,这样的数据将要从第二分配区域存储和读取)。可以理解,可以为单个特性或多个特性指定单个策略(例如,该策略指定要在何处存储热数据,要在何处存储冷数据,要在何处存储随机访问数据,将要在何处存储顺序访问数据)和/或可以为个体特性或特性对指定一个或多个策略(例如,用于热数据和冷数据的第一策略、用于用户数据和元数据的第二策略等)。可以将策略指派给要应用这样的策略的分配区域。策略可以为要应用这样的策略的分配区域指定流标识符。

[0050] 在一个示例中,第一特性可以对应于用户数据特性,并且第二特性可以对应于元数据特性(例如,元数据可以比用户数据更频繁地被覆写,并且因此具有不同的访问模式而应当被分离地存储)。在另一示例中,第一特性可以对应于第一数据频率访问特性,并且第二特性可以对应于第二数据频率访问特性(例如,更频繁地访问的数据(诸如热数据)可以存储在与不频繁地访问的数据(诸如冷数据)不同的分配区域内)。在另一示例中,第一特性可以对应于顺序访问特性,并且第二特性可以对应于随机访问特性(例如,顺序访问数据可以存储在与随机访问数据不同的分配区域内)。在另一示例中,第一特性可以对应于第一存储聚合特性,并且第二特性可以对应于第二存储聚合特性(例如,提供给第一客户端的第一存储聚合的数据可以存储在与提供给第二客户端的第二存储聚合的数据不同的分配区域内)。可以理解,可以在策略内定义各种其他特性。

[0051] 在一个示例中,针对存储设备定义了一组分配区域。可以将策略指派给该组分配区域中的分配区域。基于策略、每个分配区域的可用空闲空间量和/或其他排序标准,对该组分配区域进行动态排序(例如,在由文件系统接收写入流用于处理时即时地排序和/或重新排序)作为一组经排序的分配区域(例如,如果用户数据可以存储在第一分配区域和第五分配区域内,则具有更多可用存储空间的分配区域可以排名更高并且因此被使用)。该组策略用于基于数据的特性将数据指派给流。该组策略还用于用适当的流标识符来标记流(例如,策略可以指示元数据将被存储在第三分配区域内,并且因此该策略用于将写入流的元数据指派给流并且该策略用于用第三分配区域的流标识符来标记该流)。

[0052] 在310处,接收用以写入存储设备的写入数据流。在一个示例中,文件系统接收写入流。可以标识数据的特性,诸如用户数据、元数据和/或诸如随机访问数据的其他类型的数据。可以评估该组经排序的分配区域以标识将被用于处理用户数据和元数据的分配区

域。例如，策略可以指定第一分配区域和/或其他分配区域将被用于处理用户数据。该策略或不同策略可以指定第二分配区域和/或其他分配区域将被用于处理元数据。因此，可以基于目标分配区域具有高于关于用户数据的阈值的排序排名（例如，可以用于存储用户数据的分配区域（诸如第一分配区域）的最高排名），从该组经排序的分配区域中选择用于存储用户数据的该目标分配区域。可以基于目标分配区域具有高于关于元数据的排序排名的阈值（例如，可以用于存储元数据的分配区域（诸如第二分配区域）的最高排名），从该组经排序的分配区域中选择用于存储元数据的该目标分配区域。

[0053] 在312处，可以通过用相应分配区域的流标识符标记的流将写入流的数据提供给存储设备。例如，可以将用户数据指派给第一流。可以使用将被用于处理用户数据的第一分配区域的第一流标识符来标记第一流。以这种方式，当存储设备接收到第一流时，存储设备将基于第一流标识符而知道使用第一分配区域来处理第一流的用户数据。可以将元数据指派给第二流。可以使用将被用于处理元数据的第二分配区域的第二流标识符来标记第二流。以这种方式，当存储设备接收到第二流时，存储设备将基于第二流标识符而知道使用第二分配区域来处理第二流的元数据。

[0054] 分配区域可以被确定为具有低于阈值的空闲空间量。在一个示例中，可以终止针对该分配区域的策略（例如，自动地终止或可以向存储管理员提供用于终止策略的建议）。在另一示例中，可以重新定义分配区域以增加空闲空间量，并且可以为分配区域保留策略。

[0055] 选择性地将相似类型的数据存储在一起并且将不相似类型的数据存储在分离的分配区域中能减少碎片，减小写入放大并且提高存储效率。

[0056] 图4示出了用于使用流选择性地分配区域中存储数据的系统400的示例。可以为存储设备416（诸如为存储设备416的存储介质418）定义一组分配区域。例如，可以将第一分配区域420定义为涵盖诸如从虚拟块编号（0）到虚拟块编号（49）的第一块范围。可以将第二分配区域422定义为涵盖诸如从虚拟块编号（50）到虚拟块编号（99）的第二块范围。可以将第三分配区域424定义为涵盖诸如从虚拟块编号（100）到虚拟块编号（149）的第三块范围。可以理解，分配区域可以彼此具有相同或不同的大小，并且存储设备416的整个存储空间或存储空间的仅部分可以用于定义分配区域。以这种方式，可以为存储设备416定义任何数目的分配区域。

[0057] 可以将策略402指派给分配区域。在一个示例中，可以将第一策略404指派给第一分配区域420和/或其他分配区域。第一策略404可以指定热数据（例如，以高于阈值频率的频率而被访问的数据）将被存储在第一分配区域420和/或其他分配区域内。可以将第二策略406指派给第二分配区域422和/或其他分配区域。第二策略406可以指定冷数据（例如，以低于阈值频率的频率而被访问的数据）将被存储在第二分配区域422和/或其他分配区域内。

[0058] 可以将第三策略408指派给第三分配区域424和/或其他分配区域。第三策略408可以指定用户数据（例如，用户文本文档）将被存储在第三分配区域424和/或其他分配区域内。可以将第四策略410指派给第四分配区域和/或其他分配区域。第四策略410可以指定元数据（例如，由存储文件系统维持的元数据，诸如卷大小信息、伙伴（partner）存储控制器信息、复制策略信息、备份策略等）将被存储在第四分配区域和/或其他分配区域内。

[0059] 可以将第五策略412指派给第五分配区域和/或其他分配区域。第五策略412可以

指定随机访问数据将被存储在第五分配区域和/或其他分配区域内。可以将第六策略414指派给第六分配区域和/或其他分配区域。第六策略414可以指定顺序访问数据将被存储在第六分配区域和/或其他分配区域内。

[0060] 可以将第七策略指派给第七分配区域和/或其他分配区域。第七策略可以指定第一聚合的数据将被存储在第七分配区域和/或其他分配区域内。可以将第八策略指派给第八分配区域和/或其他分配区域。第八策略可以指定第二聚合的数据将被存储在第八分配区域和/或其他分配区域内。

[0061] 可以理解,可以将任何数目的策略指派给单个分配区域(例如,第九策略以及第十策略,第九策略指定热数据可以被存储在第七分配区域内的,第十策略指定元数据可以被存储在第七分配区域内的),并且可以将策略指派给多于一个分配区域(例如,指定可以将热数据存储在第一分配区域420、第七分配区域和第九分配区域内的策略)。还可以理解,策略可以应用于数据的单个分类(例如,热数据)或者可以应用于数据的多个分类(例如,指定要在何处存储热数据,要在何处存储冷数据,要在何处存储用户数据等等的策略)。

[0062] 图5示出了用于使用流选择性地分配区域中存储数据的系统500的示例。可以跨多个存储设备502定义分配区域。例如,可以跨第一存储设备504、第二存储设备506、第三存储设备508、奇偶校验存储设备510和/或其他存储设备(例如,具有RAID配置的存储设备)来定义第一分配区域512。以这种方式,用与第一分配区域512相关联的流标识符来标记的流内的数据可以被存储在第一分配区域512内,诸如跨第一存储设备504、第二存储设备506、第三存储设备508和/或奇偶校验存储设备510中的一个或多个存储设备而被存储。

[0063] 可以跨第一存储设备504、第二存储设备506、第三存储设备508、奇偶校验存储设备510和/或其他存储设备的第二部分来定义第二分配区域514。以这种方式,用与第二分配区域514相关联的流标识符来标记的流内的数据可以被存储在第二分配区域514内,诸如跨第一存储设备504、第二存储设备506、第三存储设备508和/或奇偶校验存储设备510中的一个或多个存储设备而被存储。

[0064] 可以跨第一存储设备504、第二存储设备506、第三存储设备508、奇偶校验存储设备510和/或其他存储设备的第三部分来定义第三分配区域516。以这种方式,用与第三分配区域516相关联的流标识符来标记的流内的数据可以被存储在第三分配区域516内,诸如跨第一存储设备504、第二存储设备506、第三存储设备508和/或奇偶校验存储设备510中的一个或多个存储设备而被存储。

[0065] 可以理解,可以为单个存储设备或跨任何数目的存储设备来定义任何数目的分配区域。还可以理解,分配区域可以限定在单个存储设备内或跨任何数目的存储设备之间来限定。

[0066] 图6的示例性方法600示出了使用流选择性地分配区域中存储数据的一个实施例。在602处,可以接收第一数据和第二数据。例如,存储控制器的文件系统可以从客户端应用接收包括第一数据和第二数据的写入流。在一个示例中,文件系统可以与诸如虚拟化层的间接层相关联,该虚拟化层虚拟化第一数据和第二数据将被写入到的存储设备的数据。

[0067] 策略可以将第一特性定义为用户数据特性,将第二特性定义为元数据特性,将第三特性定义为热数据特性,将第四特性定义为冷数据特性,等等。在604处,可以将第一数据标识为具有在策略内定义的第一特性(例如,第一数据是用户数据库中的用户数据)。在606

处,可以将第二数据标识为具有在策略内定义的第二特性(例如,第二数据是由存储控制器用以管理用户数据库的复制的元数据)。以这种方式,第一数据可以被标识为用户数据,并且第二数据可以被标识为元数据。

[0068] 在608处,将第一数据指派给第一流。第一流用由策略为用户数据的第一特性所指定的第一流标识符来标记。第一流标识符与存储设备的第一分配区域相关联,该第一分配区域是由策略定义用于存储用户数据(例如,文件系统和存储设备可以已经协商以确定将使用第一流标识符来标记将使用第一分配区域而被处理的用户数据的流)。

[0069] 在610处,将第二数据指派给第二流。第二流用由策略为元数据的第二特性所指定的第二流标识符来标记。第二流标识符与存储设备的第二分配区域相关联,该第二分配区域是由策略定义用于存储元数据(例如,文件系统和存储设备可以已经协商以确定将使用第二流标识符来标记将使用第二分配区域而被处理的元数据的流)。

[0070] 在612处,基于用第一流标识符而被标记的第一流,将第一流发送到存储设备以用于将用户数据的第一数据写入第一分配区域。以这种方式,可以将用户数据选择性地存储在第一分配区域内,而不存储在未被指派用于用户数据的其他分配区域内。

[0071] 在614处,基于用第二流标识符而被标记的第二流,将第二流发送到存储设备以用于将元数据的第二数据写入第二分配区域。以这种方式,可以将元数据选择性地存储在第二分配区域内,而不存储在未被指派用于元数据的其他分配区域内。

[0072] 图7示出了用于使用流选择性地向分配区域中存储数据的系统700的示例。文件系统702或任何其他硬件或软件模块可以在存储设备712内定义一个或多个分配区域。例如,可以为存储设备712的第一块范围(例如,虚拟块编号的第一范围)定义第一分配区域714。可以为存储设备712的第二块范围(例如,虚拟块编号的第二范围)定义第二分配区域716。可以为存储设备712的第三块范围(例如,虚拟块编号的第三范围)定义第三分配区域718。以这种方式,可以在存储设备712内和/或跨其他存储设备定义多个分配区域。

[0073] 文件系统702可以与存储设备712协商以确定文件系统702将用以标记数据流的流标识符。流标识符将是给存储设备712的、用该流标识符来标记的流的数据将要在相应的分配区域内被处理(例如,存储)的指示符。例如,可以为第一分配区域714指定第一流标识符。可以为第二分配区域716指定第二流标识符724。可以为第三分配区域718指定第三流标识符726。

[0074] 文件系统702可以将策略704指派给分配区域。例如,可以将策略指派给第二分配区域716和/或第三分配区域718。策略可以指定热数据(例如,以大于阈值的频率而被访问的数据)将被存储在第二分配区域716内,并且该热数据的流将用为第二分配区域716指定的第二流标识符724而被标记。以这种方式,热数据将被存储/包含在第二分配区域716中,而不是在其他分配区域中。因此,由于频繁访问热数据而产生的碎片可以被包含在第二分配区域716内,并且不会将附加的碎片引入到其他分配区域。

[0075] 策略可以指定冷数据(例如,以低于阈值的频率而被访问的数据)将被存储在第三分配区域718内,并且该冷数据的流将用为第三分配区域718指定的第三流标识符726而被标记。以这种方式,冷数据将被存储/包含在第三分配区域718内,而不是在其他分配区域中。因此,垃圾回收技术和/或将有效数据从目的地单元移动到空闲单元以便可以将新数据写入这些目的地单元的其他技术不会不必要地来回移动冷数据。

[0076] 可以将分配区域排序为一组经排序的分配区域。可以基于策略704、可用空闲空间和/或其他排序标准对分配区域进行排序。例如,策略可以指定可以将随机访问数据存储在第二分配区域716和第三分配区域718内。当接收到随机访问数据的写入流时,可以关于将随机访问数据存储在存储设备712内的当前场景来动态地对分配区域排序。基于指定第二分配区域716和第三分配区域718而不是第一分配区域714将被用于存储随机访问数据的策略,第一分配区域714可以被排名为低于第二分配区域716和第三分配区域718。基于第二分配区域716具有多于第三分配区域718的可用空闲空间,可以将第二分配区域716排名为高于第三分配区域718。

[0077] 文件系统702可以接收写入流706。写入流706可以包括热数据708(例如,以大于阈值的频率而被访问的数据)、冷数据710(例如,以低于阈值的频率而被访问的数据)和/或其他特性的数据。文件系统702可以利用策略704来确定(例如,对分配区域排序并且选择排名最高的分配区域)第二分配区域716将用于存储热数据708。文件系统702可以将热数据708指派给第一流720。文件系统702可以用第二分配区域716的第二流标识符724来标记第一流720。以这种方式,第一流720被提供给存储设备712。基于第一流720使用第二流标识符724被标记,存储设备712将使用第二分配区域716来处理第一流720的热数据708。

[0078] 文件系统702可以利用策略704来确定(例如,对分配区域排序并且选择排名最高的分配区域)第三分配区域718将用于存储冷数据710。文件系统702可以将冷数据710指派给第二流722。文件系统702可以用第三分配区域718的第三流标识符726标记来第二流722。以这种方式,第二流722被提供给存储设备712。基于第二流722使用第三流标识符726被标记,存储设备712将使用第三分配区域718来处理第二流722的冷数据710。

[0079] 又一实施例涉及计算机可读介质,其包括被配置为实现本文中呈现的一种或多种技术的处理器可执行指令。以这些方式设计的计算机可读介质或计算机可读设备的示例实施例在图8中示出,其中实现800包括在其上编码有计算机可读数据806的计算机可读介质808,诸如可记录光盘(CD-R)、可记录数字多功能光盘(DVD-R)、闪存驱动器、硬盘驱动器盘等。该计算机可读数据806(诸如包括0或1中的至少一个的二进制数据)依次包括被配置为根据本文中阐述的一个或多个原理而操作的处理器可执行计算机指令804。在一些实施例中,例如,处理器可执行计算机指令804被配置为执行方法802,诸如图3的示例性方法300中的至少一些方法和/或图6的示例性方法600中的至少一些方法。在一些实施例中,例如,处理器可执行计算机指令804被配置为实现系统,诸如图4的示例性系统400中的至少一些系统、图5的示例性系统500中的至少一些和/或图7的示例性系统700中的至少一些系统。许多这样的计算机可读介质被预期以根据本文中给出的技术来操作。

[0080] 应当理解,本文中描述的过程、架构和/或过程可以以硬件、固件和/或软件来实现。还应当理解,本文中阐述的规定可以适用于实施为或包括存储系统的任何类型的专用计算机(例如,文件主机、存储服务器和/或存储服务设备)和/或通用计算机,其包括独立计算机或其部分。此外,本文中的教导可以被配置为各种存储系统架构,包括但不限于网络附属存储环境和/或直接附接到客户端或主计算机的存储区域网络和盘组件。因此,存储系统应当被广泛地理解为除了包括被配置为执行存储功能并且与其他设备或系统相关联的任何子系统之外,还包括这样的布置。

[0081] 在一些实施例中,本公开中描述和/或示出的方法可以全部或部分地在计算机可

读介质上实现。计算机可读介质可以包括被配置为实现本文中给出的一个或多个方法的处理器可执行指令,并且可以包括用于存储其后可以由计算机系统读取的数据的任何机制。计算机可读介质的示例包括(硬)驱动器(例如,经由网络附属存储(NAS)可访问)、存储区域网络(SAN)、易失性和非易失性存储器(诸如只读存储器(ROM)、随机存取存储器(RAM)、电可擦除可编程只读存储器(EEPROM)和/或闪存)、光盘只读存储器(CD-ROM)、CD-R、可重写光盘(CD-RW)、DVD、盒式磁带、磁带、磁盘存储器、光学或非光学数据存储设备和/或可以用于存储数据的任何其他介质。

[0082] 尽管用特定于结构特征或方法动作的语言描述了本主题,但是应当理解,所附权利要求书中限定的主题不必限于上述特定的特征或动作。相反,上面描述的特定的特征和动作是作为实现至少一些权利要求的示例形式而公开的。

[0083] 本文中提供了实施例的各种操作。描述一些或所有操作的顺序不应当被解释为暗示这些操作必须依赖于该顺序。考虑到本说明书的益处,可以理解备选的顺序。此外,应当理解,并非所有操作都必须存在于本文中提供的每个实施例中。而且,应当理解,在一些实施例中并非所有操作都是必需的。

[0084] 此外,所要求保护的主体被实现为使用标准应用或工程技术来生产软件、固件、硬件或其任何组合以控制计算机实现所公开的主题的方法、装置或制品。本文中使用的术语“制品”旨在涵盖从任何计算机可读设备、载体或介质可访问的计算机应用。当然,在不脱离所要求保护的主体范围或精神的情况下,可以对该配置进行很多修改。

[0085] 如本申请中使用的,术语“组件”、“模块”、“系统”、“接口”等通常旨在指代与计算机相关的实体,其是硬件、硬件和软件的组合、软件、或执行软件。例如,组件包括在处理器上运行的进程、处理器、对象、可执行文件、执行线程、应用或计算机。作为示例,在控制器上运行的应用和控制器都可以是组件。驻留在进程或执行线程中的一个或多个组件和组件可以位于一台计算机上或者分布在两个或多个计算机之间。

[0086] 此外,在本文中“示例性”意味着用作示例、实例、说明等,并且不一定是作为有利的。如本申请中使用的,“或”旨在表示包含性的“或”而非排他性的“或”。另外,除非另有说明或从上下文清楚地指向单数形式,否则本申请中使用的“一”和“一个”通常被解释为意指“一个或多个”。并且,A和B的至少一个和/或类似表述通常表示A或B和/或A和B两者。此外,对于“包含”、“具有”、“有”、“涵盖”或其变型的使用的范围,这样的术语旨在是包括性的,其与术语“包括”类似。

[0087] 在不脱离所要求保护的主体范围或精神的情况下,可以对本公开做出很多修改。除非另有说明,否则“第一”、“第二”等不旨在暗示时间方面、空间方面、排序等。相反,这样的术语仅用作针对特征、要素、项目等的标识符、名称等。例如,第一组信息和第二组信息通常对应于信息组A和信息组B,或两组不同或两组等同的信息或同一组信息。

[0088] 此外,尽管已经关于一个或多个实现示出并且描述了本公开,但是基于对本说明书和附图的阅读和理解,本领域技术人员将想到等同的改变和修改。本公开包括所有这样的修改和变更,并且仅由所附权利要求的范围所限制。特别地,关于由所描述组件(例如,元件、资源等)执行的各种功能,除非另有指示,否则被用于描述这样的组件的术语旨在对应于执行所述组件的指定功能的、尽管在结构上不等同于所公开的结构任何组件(例如,在功能上等同)。另外,尽管可能仅关于若干实现中的一个公开了本公开的特定特征,但是这

样的特征可以与其他实现的一个或多个其他特征组合,因为其是可以期望以及对于任何给定或特定应用可以是有利的。

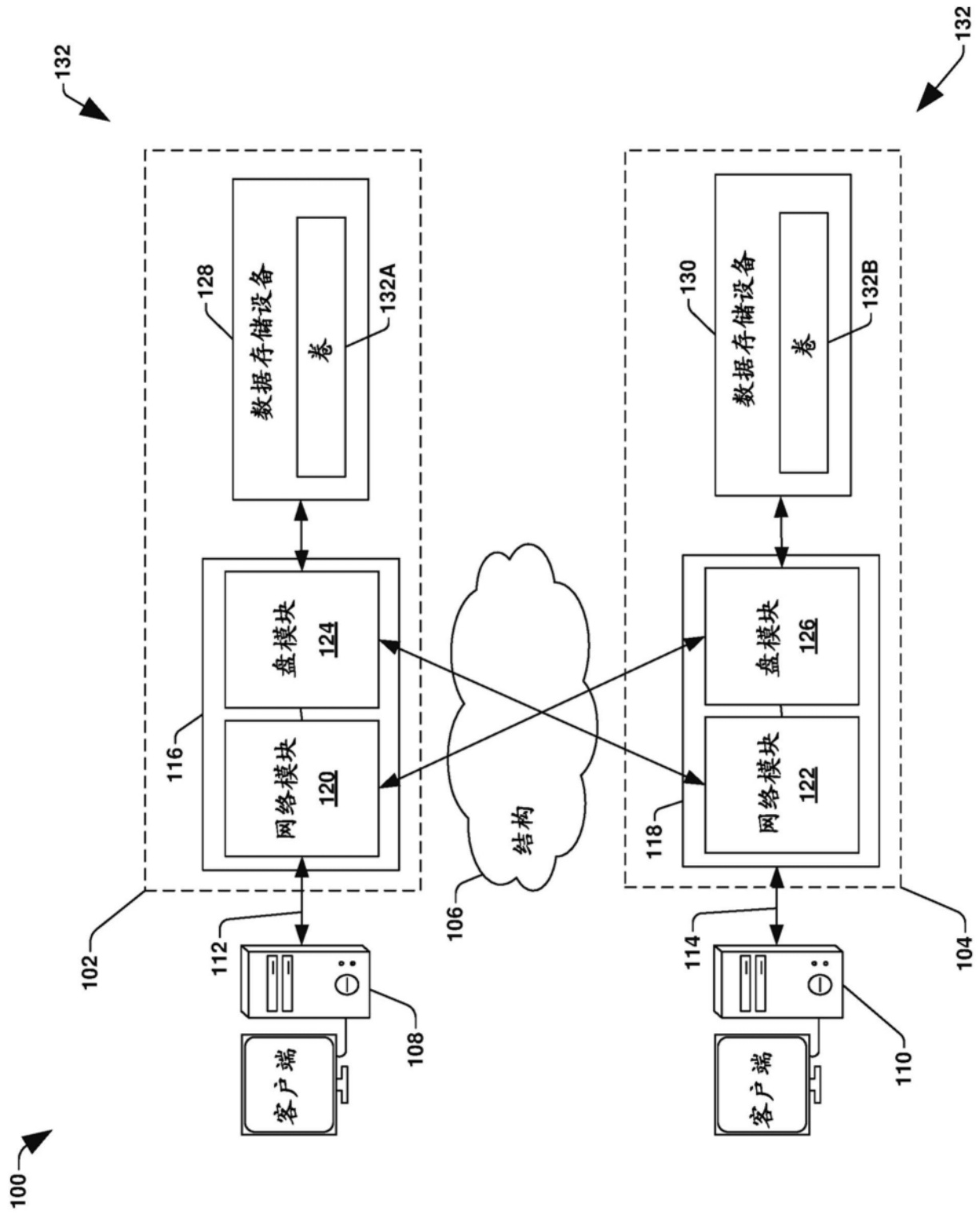


图1

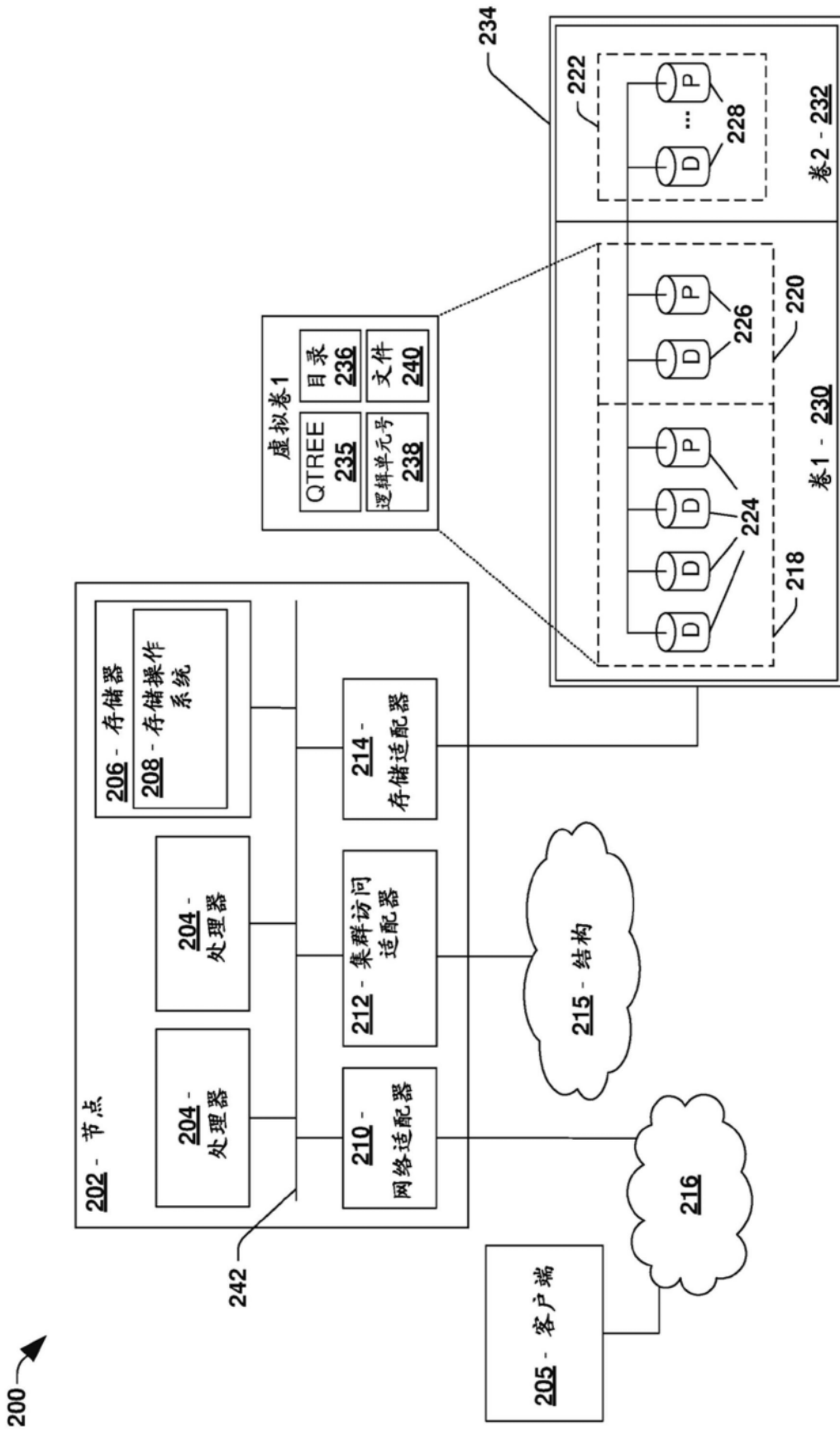


图2

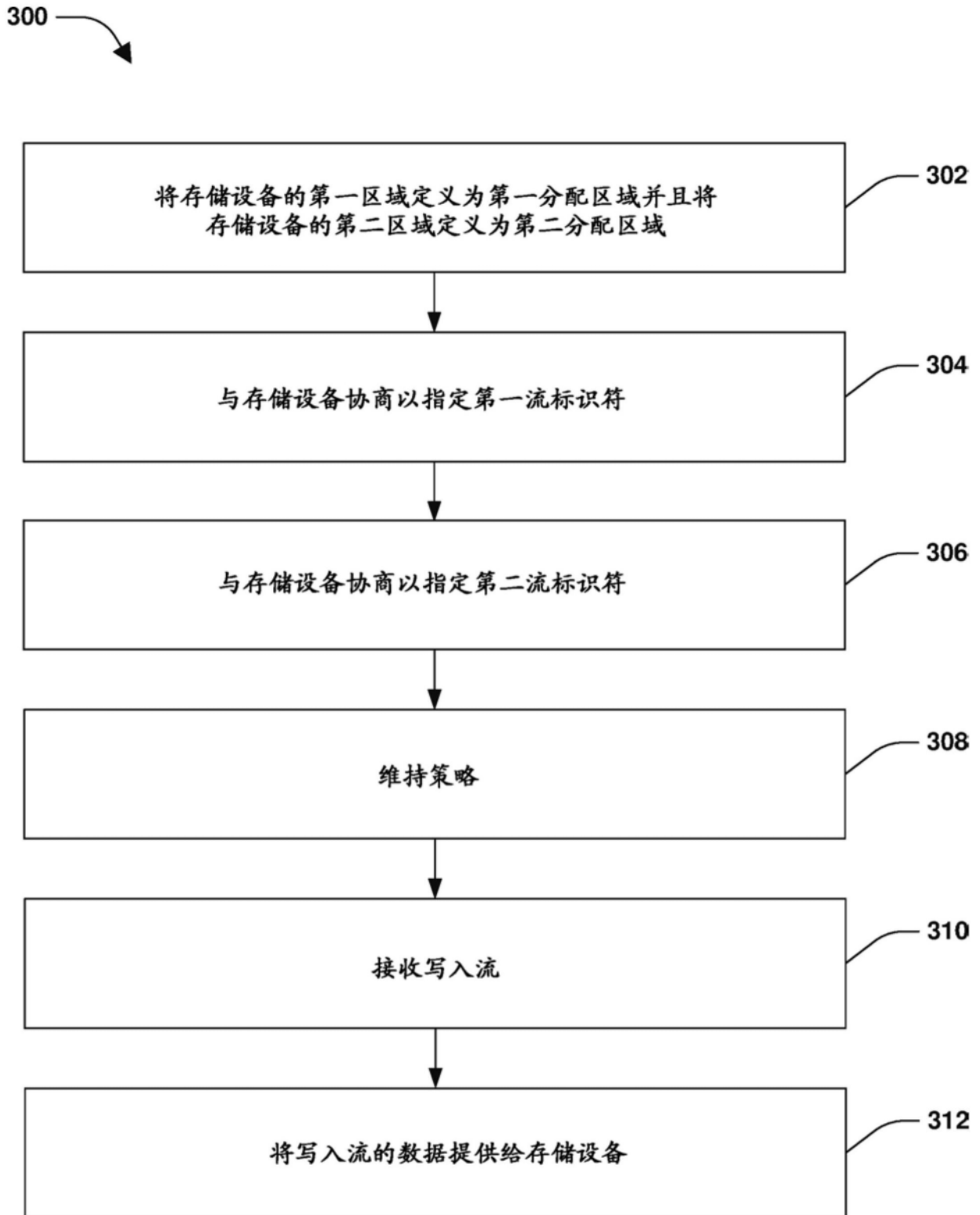


图3

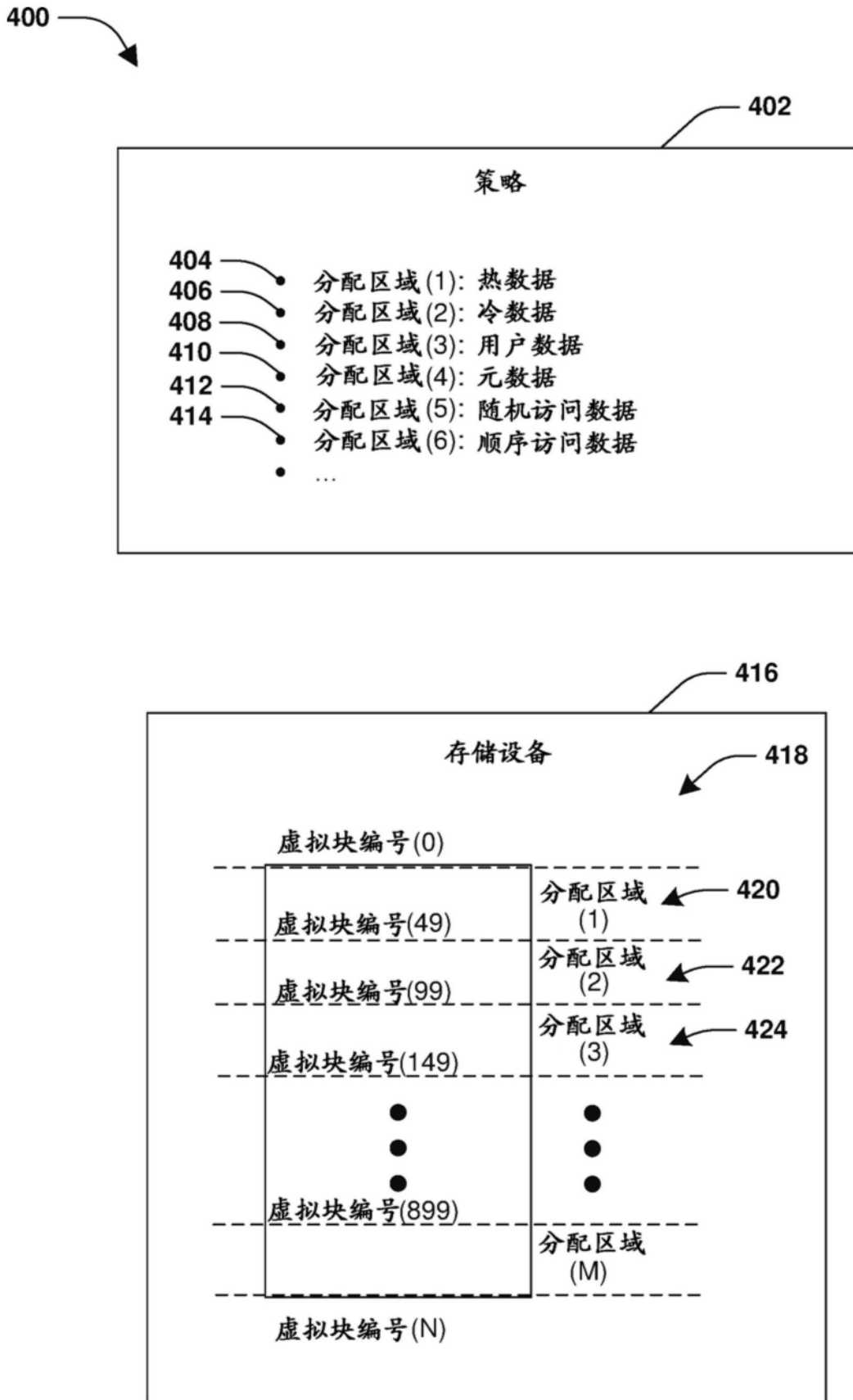


图4

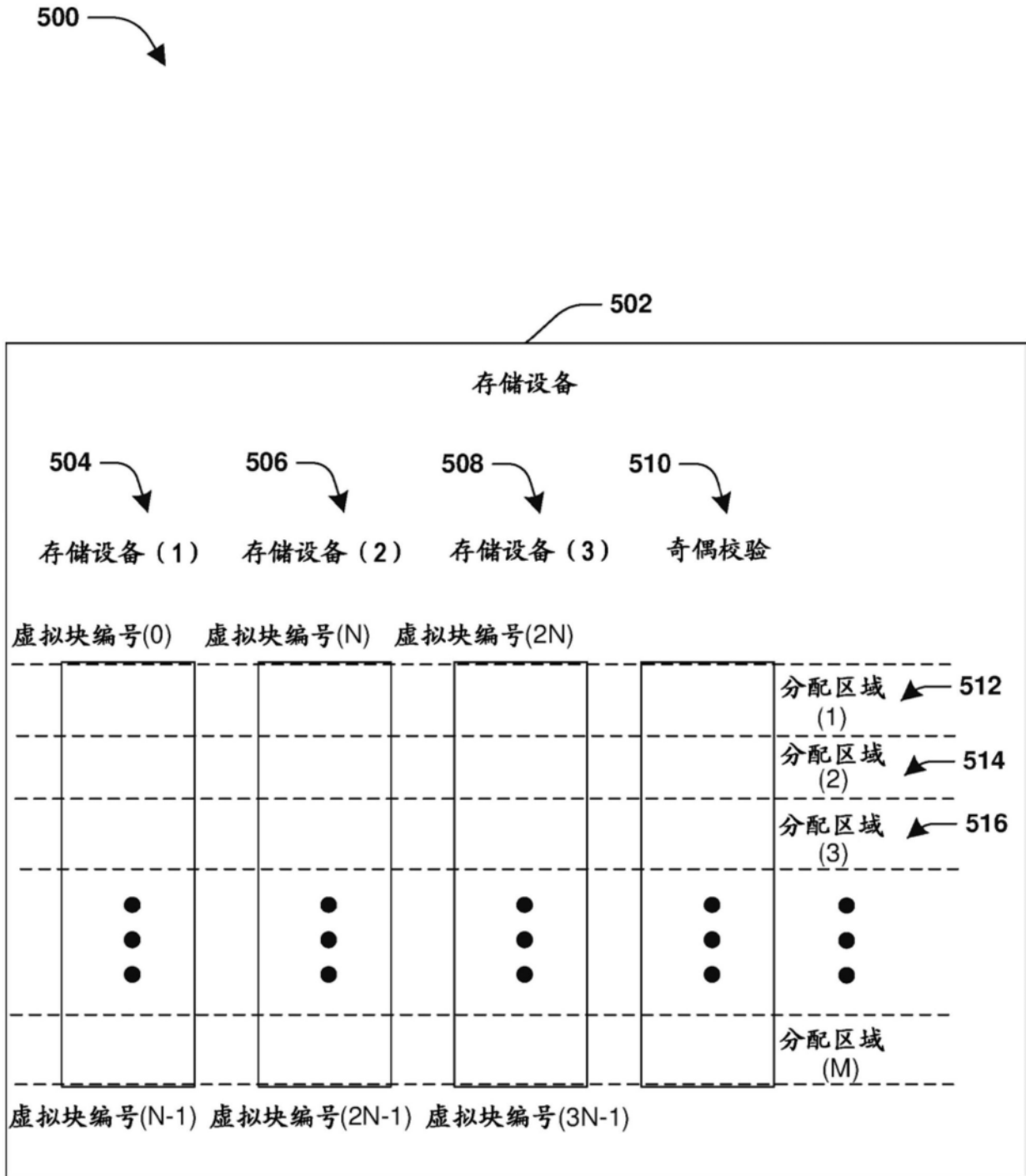


图5

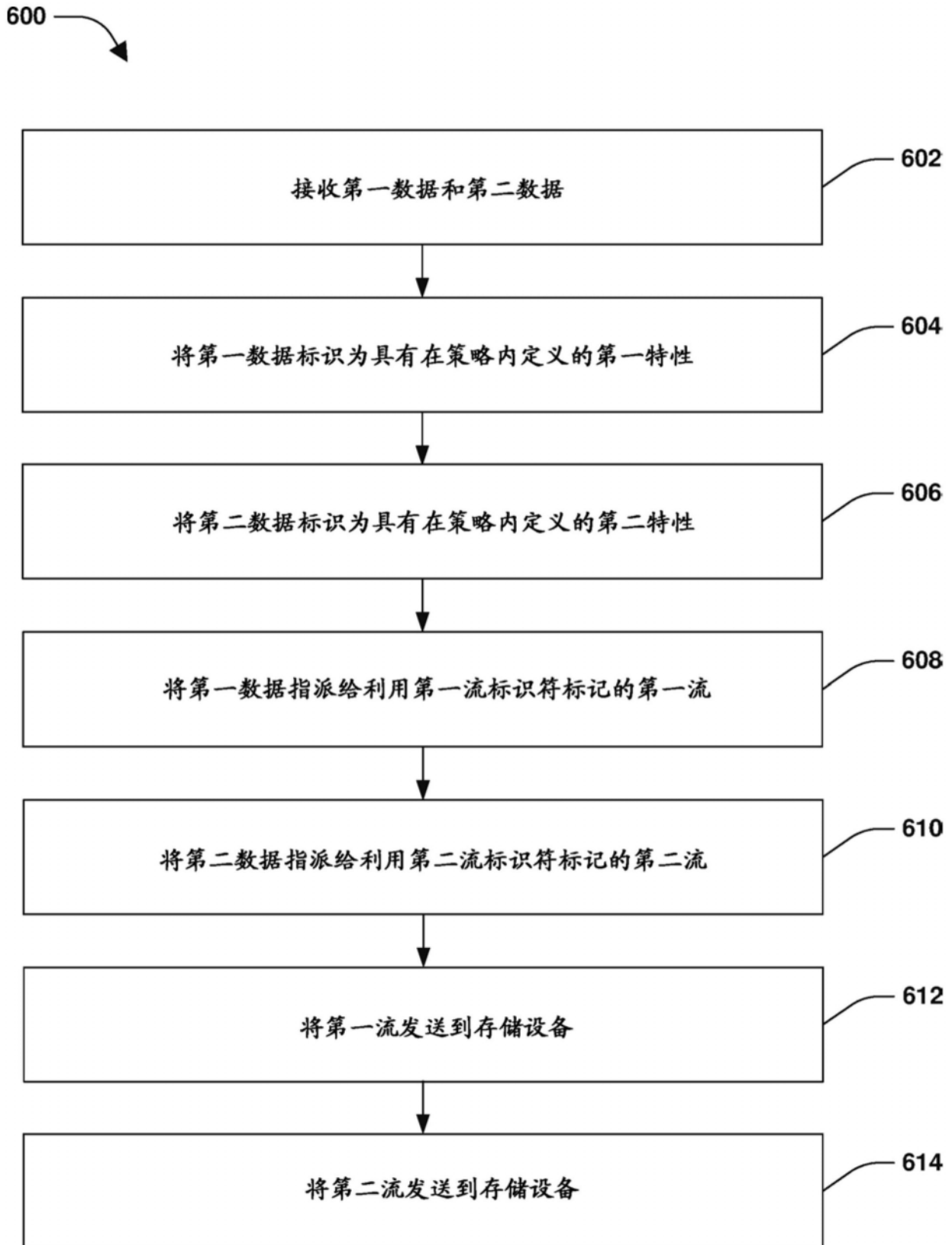


图6

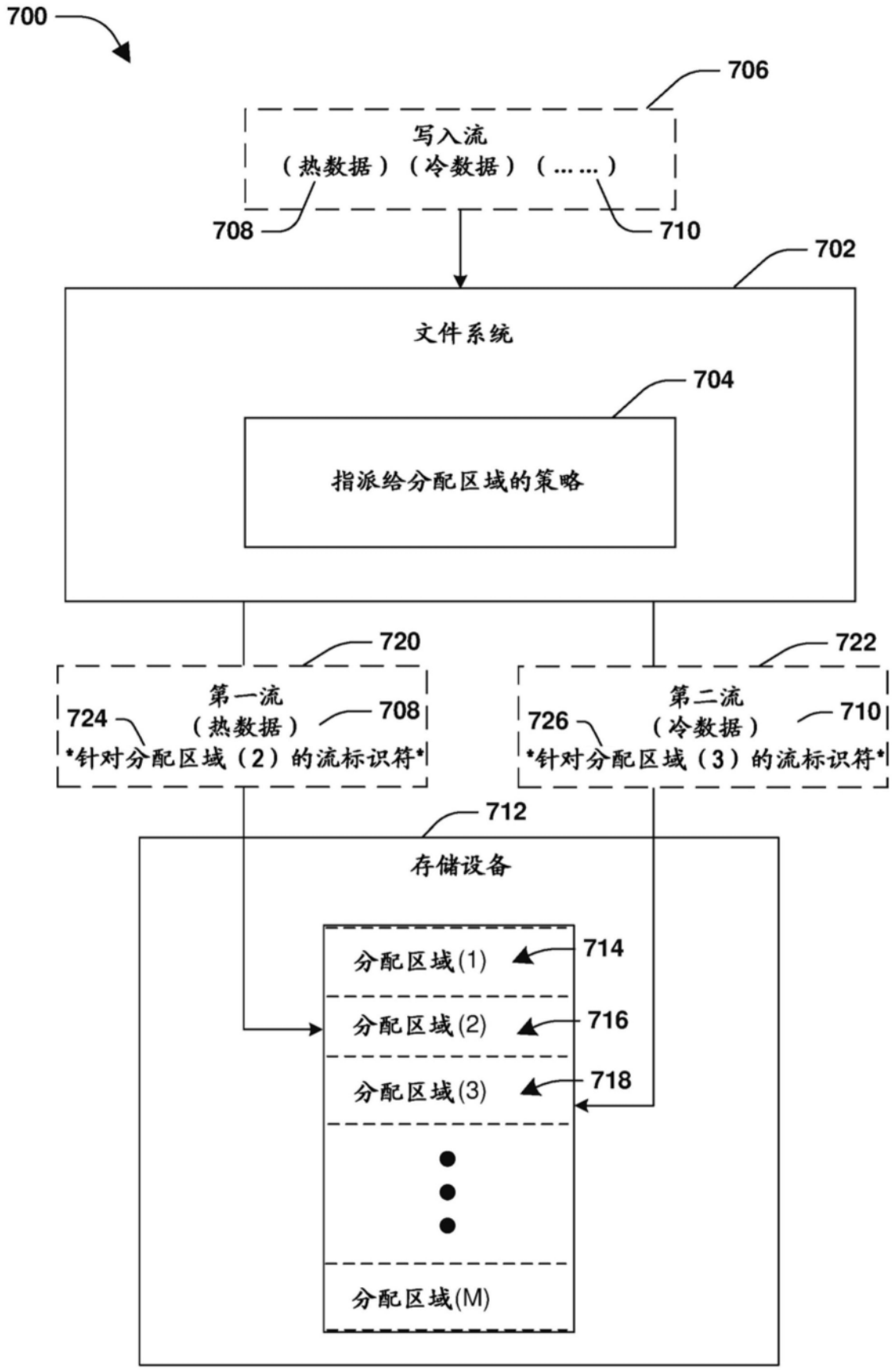


图7

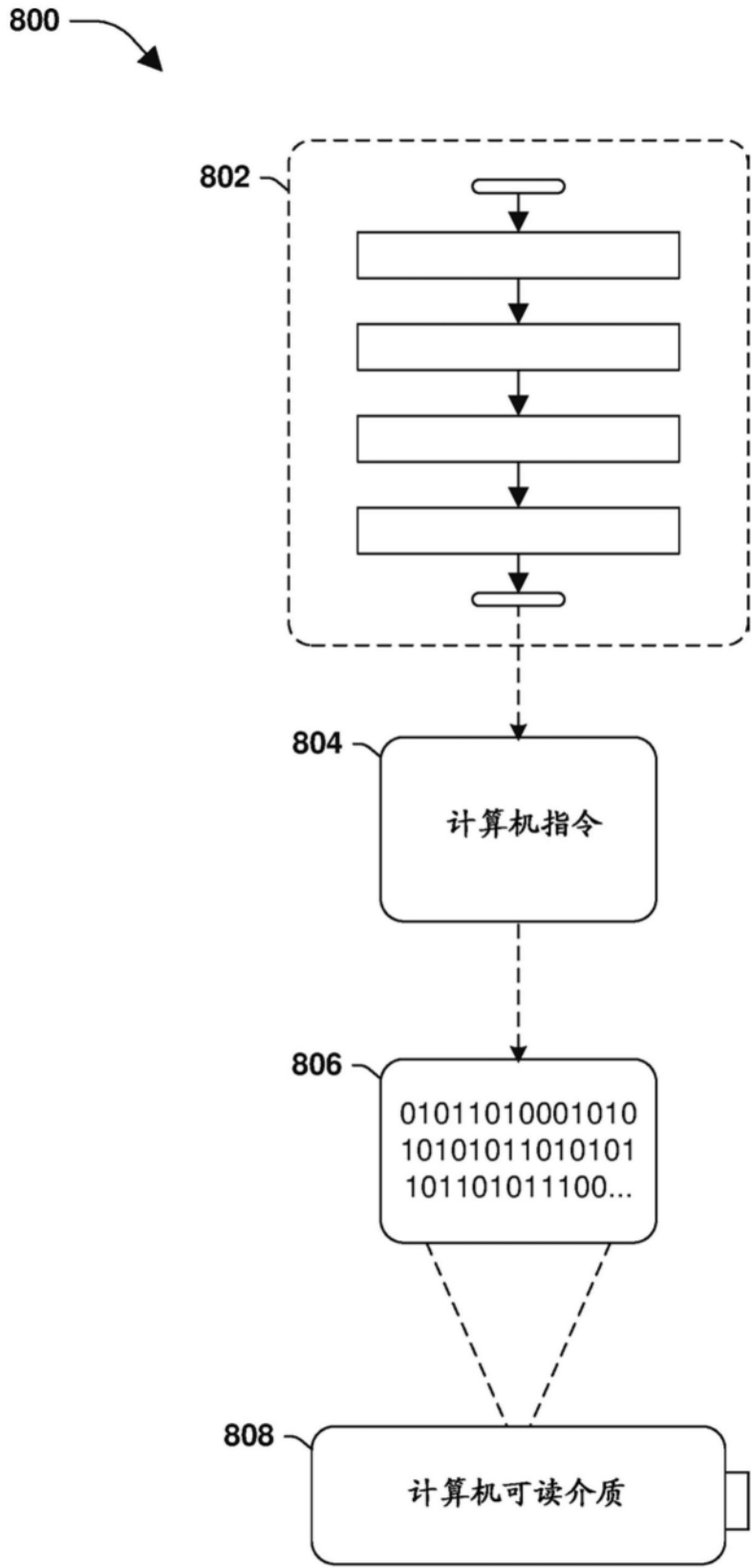


图8