



(10)授权公告号 CN 105190526 B

(45)授权公告日 2018.03.30

(21)申请号 201480008161.3

(22)申请日 2014.02.06

(65)同一申请的已公布的文献号

申请公布号 CN 105190526 A

(43)申请公布日 2015.12.23

(30)优先权数据

13/763491 2013.02.08 US

(85)PCT国际申请进入国家阶段日

2015.08.10

(86)PCT国际申请的申请数据

PCT/US2014/014971 2014.02.06

(87)PCT国际申请的公布数据

W02014/124064 EN 2014.08.14

(73)专利权人 微软技术许可有限责任公司

地址 美国华盛顿州

(72)发明人 S.卡拉莫夫 D.M.卡拉汉

(74)专利代理机构 中国专利代理(香港)有限公司 72001

代理人 刘靖龙 景军平

(51)Int.Cl.

G06F 3/06(2006.01)

G06F 3/01(2006.01)

(56)对比文件

US 2010312983 A1, 2010.12.09,

US 2011238946 A1, 2011.09.29,

CN 102598020 A, 2012.07.18,

TW 201022937 A, 2010.06.16,

审查员 陈恺

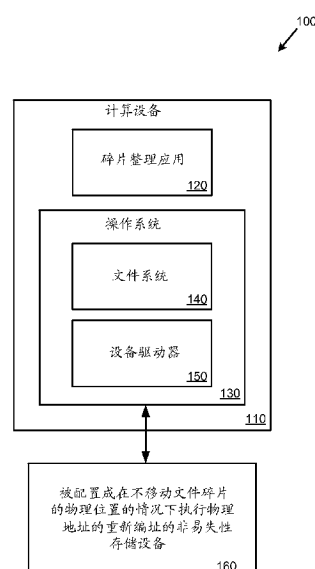
权利要求书2页 说明书12页 附图8页

(54)发明名称

对用于非易失性存储设备的存储器重新编址

(57)摘要

可以将非易失性存储设备上的用于已破碎文件的存储器重新编址到相邻物理存储器地址,同时存储在非易失性存储设备上的已破碎文件的文件碎片的物理位置在存储器被重新编址之后保持相同。可以基于重新编址的相邻物理存储器地址来更新逻辑块寻址(LBA)映射表。



1. 一种至少部分地由计算设备执行以用于执行在非易失性存储设备上的用于已破碎文件的存储器的重新编址的方法,包括:

由计算设备向非易失性存储设备发送用以对已破碎文件的存储器重新编址的命令,其中所述非易失性存储设备是固态驱动,其中,已破碎文件的文件碎片跨多个非相邻物理地址散布,并且被存储在非易失性存储设备内的多个物理位置处;以及

由计算设备从非易失性存储设备接收已破碎文件的存储器已被重新编址的响应,其中,所述存储器已被重新编址到相邻物理地址:

对于已破碎文件的起始块,分配唯一的起始物理存储器地址,其中唯一的起始物理存储器地址仅被分配给已破碎文件;

对于已破碎文件的所有后续块,分配对于已破碎文件不是唯一的可共享的物理存储器地址,其中可共享的物理存储器地址由存储在非易失性存储设备上的其它文件可共享,使得相同的给定可共享的物理存储器地址可分配给多个文件;

其中已破碎文件仅使用用于起始块的唯一的起始物理存储器地址以及已破碎文件的长度作为顺序访问被检索,其中已破碎文件的所有后续块在起始块的唯一的起始物理存储器地址之后被连续地访问;

其中,在已破碎文件的存储器已被重新编址之后,文件碎片的所述多个物理位置保持相同。

2. 一种非易失性存储设备,包括:

处理单元;以及

非易失性存储器;

所述非易失性存储设备被配置成执行用于对用于已破碎文件的存储器重新编址的操作,该操作包括:

接收用以对已破碎文件的存储器重新编址的命令,其中,已破碎文件的文件碎片跨多个非相邻物理地址散布,并且被存储在非易失性存储设备内的多个物理位置处;以及

针对文件碎片中的每个,向文件碎片分配相邻物理存储器地址,包括:

对于已破碎文件的起始块,分配唯一的起始物理存储器地址,其中唯一的起始物理存储器地址仅被分配给已破碎文件;

对于已破碎文件的所有后续块,分配对于已破碎文件不是唯一的可共享的物理存储器地址,其中可共享的物理存储器地址由存储在非易失性存储设备上的其它文件可共享,使得相同的给定可共享的物理存储器地址可分配给多个文件;

其中已破碎文件仅使用用于起始块的唯一的起始物理存储器地址以及已破碎文件的长度作为顺序访问被检索,其中已破碎文件的所有后续块在起始块的唯一的起始物理存储器地址之后被连续地访问;

其中,在已破碎文件的存储器已被重新编址之后,文件碎片的所述多个物理位置保持相同。

3. 根据权利要求2的非易失性存储设备,其中,所述非易失性存储设备是固态驱动。

4. 根据权利要求2的非易失性存储设备,其中,所述非易失性存储设备是相变存储器设备。

5. 根据权利要求2的非易失性存储设备,其中,接收用以对已破碎文件的存储器重新编

址的命令包括：

基于破碎的程度来确定要重新编址的最可能的候选文件；以及
选择该最可能的候选文件作为已破碎文件。

6. 一种存储计算机可执行指令的计算机可读存储介质，所述计算机可执行指令用于促使计算设备执行用于对用于已破碎文件的存储器重新编址的操作，该操作包括：

向非易失性存储设备发送用以对已破碎文件的存储器重新编址的命令，其中，已破碎文件的文件碎片跨多个非相邻物理地址散布，并且被存储在非易失性存储设备内的多个物理位置处；以及

从非易失性存储设备接收已破碎文件的存储器已被重新编址的响应，其中，所述存储器已被重新编址到相邻物理地址，其中所述重新编址包括：

对于已破碎文件的起始块，分配唯一的起始物理存储器地址，其中唯一的起始物理存储器地址仅被分配给已破碎文件；

对于已破碎文件的所有后续块，分配对于已破碎文件不是唯一的可共享的物理存储器地址，其中可共享的物理存储器地址由存储在非易失性存储设备上的其它文件可共享，使得相同的给定可共享的物理存储器地址可分配给多个文件；

其中已破碎文件仅使用用于起始块的唯一的起始物理存储器地址以及已破碎文件的长度作为顺序访问被检索，其中已破碎文件的所有后续块在起始块的唯一的起始物理存储器地址之后被连续地访问；

其中，在已破碎文件的存储器已被重新编址之后，文件碎片的所述多个物理位置保持相同。

对用于非易失性存储设备的存储器重新编址

背景技术

[0001] 随着反复地在存储设备上写入和擦除文件,随时间推移这些文件可能变得破碎,从而降低存储设备的性能。为了帮助减轻此性能问题,可在存储设备上执行磁盘碎片整理。磁盘碎片整理指的是这样的操作,其通过将存储设备上的文件碎片移动至相邻(contiguous)位置来减少存储设备上的文件破碎,从而减少读入或写出所有文件碎片所需的存储设备与中央处理单元(CPU)存储器之间的输入/输出(I/O)事务数目。

[0002] 替代或者附加于传统硬盘驱动(诸如旋转磁和光驱),诸如固态驱动(SSD)之类的非易失性存储设备已经逐渐地被用作存储设备。虽然在传统硬盘驱动的情况下可以有效地使用碎片整理,但在非易失性存储设备的情况下使用碎片整理可能有问题,因为这些非易失性存储设备可能由于对设备的重复擦除操作而遭受磨损。由于非易失性存储设备在其可靠性受到危害之前具有其可以被擦除和写入的有限次数,所以非易失性存储设备的磁盘碎片整理遭受磁盘性能对存储设备的寿命的权衡。

发明内容

[0003] 提供本发明内容是为了以简化形式来介绍下面在详细描述中进一步描述的概念的选择。本发明内容并不意图识别要求保护的主题的关键特征或必要特征,其也不意图用来限制要求保护的主题的范围。

[0004] 描述了用于重新布置在非易失性存储设备中的存储器地址的技术和工具。例如,可以在不将数据从其在存储设备上的物理位置移动的情况下对存储器地址进行重新编址。存储设备可以以对操作系统而言透明的方式对存储器地址进行重新编址。替换地,操作系统可向存储设备发布命令以执行优化并针对已优化存储设备修改例如映射表。

[0005] 例如,可以提供一种用于执行在非易失性存储设备上的用于已破碎文件的存储器的重新编址的方法。该方法包括向非易失性存储设备发送用以对已破碎文件的存储器重新编址的命令,其中,已破碎文件的文件碎片跨多个非相邻物理地址散布,并且从非易失性存储设备接收用于已破碎文件的存储器已被重新编址到相邻物理地址的响应。文件碎片的物理位置在存储器已被重新编址之后保持相同。

[0006] 作为另一示例,可以将非易失性存储设备配置成执行本文所描述的操作。例如,非易失性存储设备可以接收用以对已破碎文件的存储器重新编址的命令,以及针对已破碎文件的文件碎片中的每个向文件碎片分配相邻物理存储器地址。文件碎片的物理位置在存储器已被重新编址之后保持相同。

[0007] 作为又一示例,可以提供一种用于促使系统执行本文所描述操作的存储计算机可执行指令的计算机可读存储介质。例如,计算机可读存储介质可以从非易失性存储设备接收用于已破碎文件的存储器已被重新编址到相邻物理地址的响应,并且可以基于重新编址的相邻物理地址来更新虚拟映射表。文件碎片的物理位置在存储器已被重新编址之后保持相同。用于操作系统的逻辑块寻址(LBA)映射表并未基于重新编址的物理地址被更新并且LBA映射表与虚拟映射表通信。

[0008] 如本文所描述,可以根据需要将多种其它特征和优点结合到技术中。

附图说明

[0009] 图1是示例性操作环境的框图。

[0010] 图2是用于执行存储器的重新编址的示例性方法的流程图。

[0011] 图3是用于执行存储器的重新编址的示例性方法的流程图。

[0012] 图4a、4b和4c是示出了对物理地址重新编址的示例的图。

[0013] 图5a和5b是示出了在不移动存储器的物理位置的同时对物理地址重新编址的示例的图。

[0014] 图6a和6b是示出了LBA映射表和物理地址的映射的示例的表格。

[0015] 图7是其中可以实现某些所描述实施例的示例性计算系统的图。

[0016] 图8是可以结合本文所描述的技术使用的示例性移动设备。

具体实施方式

[0017] 示例1—示例性概述

[0018] 以下描述针对用于对在非易失性存储设备上的物理存储器地址重新编址的技术和解决方案。例如,可以在不将存储器从其在存储设备上的物理存储器位置移动的情况下对已破碎文件的存储器的物理地址重新编址。

[0019] 通过对存储器地址重新编址,可以将已破碎文件的文件碎片重新编址到相邻存储器地址,从而允许进行更高效的文件操作(例如,文件检索)。例如,如果文件的文件碎片位于相邻存储器地址处,操作系统可能能够向非易失性存储设备作出单个请求或打包多个请求以检索文件。另一方面,如果文件位于非相邻存储器地址处,则操作系统可能必须向存储设备作出多个请求以检索文件。

[0020] 非易失性存储设备的磁盘碎片整理将潜在地实现类似的效果。通过对存储设备进行碎片整理,将在存储设备上的实际物理存储器位置之间移动文件碎片,使得文件碎片在碎片整理之后将位于相邻物理存储器位置处。然而,碎片整理可能缩短诸如SSD之类的非易失性存储设备的使用寿命,因为每个碎片整理操作将要求多次擦除和写入操作以在存储设备中到处移动文件碎片,从而增加了存储设备上的磨损。由于在行业中称为写入放大的现象,附加磨损的问题比仅仅擦除和写入数据更大。写入放大描述了其中在向存储器重新写入之前必须将其擦除的情形。数据通常是以例如在尺寸方面为4-8千字节的页面尺寸写入的,而要擦除的块(擦除块)通常在尺寸方面大得多(例如,在某些高密度存储设备上128千字节或者甚至几MB)。因此,当写入或移动数据时,即使例如要写入512字节的数据,此写入也可导致必须移动并擦除大得多的数据块。应认识到的是,当不知道底层含意时,对这些驱动进行碎片整理可以缩短存储设备可以可靠地操作的时间,并且本文所描述的实施例可以在仍然周期性地移除破碎的同时延长存储设备可以可靠地操作的时间。

[0021] 在旋转磁或光驱的情况下的碎片整理要求将文件碎片在物理上移动至驱动上的新的邻近位置,以实现当读写头在其它文件碎片的物理附近时发生的I/O管线中的优化。本文所描述的实施例将显示操作系统如何利用非易失性存储设备在不必实际上将内容拷贝到新的物理邻近位置的情况下通过修改其中存储内容的可寻址位置来优化I/O模式。

非易失性存储设备将内容存储在可寻址位置处,所述可寻址位置可以通过将其中存储了相关内容的不同位置的查找地址修改成在逻辑上邻近被优化。本文所描述的实施例还在不招致存储设备上的过早磨损的损伤效应的情况下向碎片整理提供类似的I/O性能优点,并且避免了与重新布置大量的存储系统内容而不是最终用户任务(诸如保存照片或播放电影)相关联的耗电功率和最终用户影响。

[0022] 示例2—示例性非易失性存储设备

[0023] 如本文所使用的,非易失性存储设备指的是在不要求通电的情况下保持其信息的任何基于半导体的存储设备。例如,非易失性存储设备可以是固态驱动、USB闪速驱动、芯片上的嵌入式存储器、相变存储器设备或任何其它类型的非易失性基于半导体的存储器。本文所描述的实施例还可以被使用在其中有序信息可能由于破碎而变成分布式的任何情形中,诸如随机存取存储器(RAM),其使用本文所描述的机制通过块或页面重新编址将各块重新排序成顺序布局而不必实际上将数据拷贝到不同的存储页面。

[0024] 如本文所使用的,非易失性存储器指的是基于半导体的存储器,并且因此不包括磁存储设备(例如,硬盘驱动)或光存储设备(例如,CD或DVD介质)。

[0025] 示例3—对物理存储器重新编址

[0026] 与磁或光存储设备相反,非易失性存储设备并不线性地读出数据。例如,在磁存储设备中,读写头移动至盘片上的位置,并且随着盘片旋转从该盘片读取信息。如果磁存储设备想要读取盘片上的另一位置处的数据,则读写头必须移动至该新位置。基于(一个或多个)盘片上的位置来布置磁存储设备的物理地址。

[0027] 另一方面,非易失性存储设备并不使用读写头,并且替代地可以通过确定单独晶体管的状态来读取信息。随着电压流过晶体管,电流被检测为二进制数据。此操作可以在许多不同的晶体管处并行地执行。虽然这些设备并未遭受与将物理读/写头移动至特定位置相关联的等待时间,但与当使用许多较小事务时相比,当操作系统和应用进行更少但更大的访问以检索或存储数据时其确实展示出性能益处。例如,从性能和功率消耗的观点出发,读取映射到一个相邻顺序文件读取请求的1 MB大块比执行每个512字节的2,000次访问以检索相同的文件有效负荷更好。采用本文所描述实施例的系统可以通过将数据转储到不同的一组块而不是释放相邻块来实现高写入速度,因为数据以如同其实际上位于物理上邻近的可寻址块中一样被寻址而结束。

[0028] 然而,使用非易失性存储设备的计算设备以及存储设备本身通常以与磁存储设备类似的方式对待非易失性存储设备,即如同其必须被以线性方式读取一样。闪存转换层(FTL)允许数据出现在特定物理位置中,并且FTL跟踪物理存储器地址到非易失性存储设备上的物理位置的映射。因此,非易失性存储设备向晶体管位置分配物理地址,使得其看起来被线性地读取。

[0029] 然而,不存在用于存储设备中的特定晶体管的物理地址和物理位置的公共映射方案,即物理地址可以映射到存储设备上的任何位置,并且附近物理地址不必映射到附近的物理位置,并且替代地,每个半导体存储设备制造商可提出其自己的方案以向存储设备分配物理存储器地址。例如,某些制造商可对存储设备中的物理存储器地址进行硬编址,而其它的制造商可动态地分配存储器地址以用于例如磨损调平。

[0030] 示例4—示例性操作环境

[0031] 在本文中的示例的任一个中,可以提供用于对存储器地址重新编址的操作环境100。图1是描绘示例性操作环境100的图。示例性操作环境100包括计算设备110,其包括碎片整理应用120和操作系统130。例如,计算设备110可以是移动计算设备,诸如移动电话或平板电脑。

[0032] 操作系统130与非易失性存储设备160进行通信。操作系统130包括文件系统140和设备驱动器150。文件系统140维持非易失性存储设备160上的文件的位置并管理对非易失性存储设备160的访问。例如,文件系统140可以是NTFS(新技术文件系统),由微软公司针对其Windows操作系统开发的文件系统。设备驱动器150控制非易失性存储设备并处理操作系统130与非易失性存储设备160之间的通信。

[0033] 在图1中,出于说明性目的将计算设备110和非易失性存储设备160示为单独部件。然而,应理解的是,计算设备110和非易失性存储设备160可以是同一设备。

[0034] 在图1中,操作系统130包含文件系统140和设备驱动器150以与非易失性存储设备160通信。然而,操作系统130可包含与非易失性存储设备160通信的其它部件。在一个实施例中,对存储器重新编址的命令可来自这些其它操作系统部件中的一个。

[0035] 在示例中,计算设备110可包含碎片整理应用120。虽然在图1中将碎片整理应用120示为在操作系统130之外,但应认识到的是可修改碎片整理应用120,使得其被集成到文件系统140中或者包括在设备驱动器150中。此外,在某些实施例中,可将碎片整理集成到非易失性存储设备160本身中。

[0036] 当在计算设备110上执行碎片整理应用120时,其可命令非易失性存储设备160对存储器地址重新编址以实现存储设备的重新编址。碎片整理应用120可以检查如何将存储在文件系统140中的每个文件通过设备驱动器150映射到非易失性存储设备160中的存储地址。当碎片整理应用120确定跨超过可配置或预定义数目的碎片(1、2、20等)存储文件时,其可以调用重新编址方法。在其它实施例中,碎片整理应用120可以使用诸如被频繁访问的文件或任何数目的其它试探法(诸如文件尺寸、系统文件、用户文件等)之类的准则。碎片整理应用120可以通过文件系统140和设备驱动器150向存储设备160发布具有已破碎的文件地址位置的命令,并且接收回具有(一个或多个)新的未破碎(或较少破碎)地址位置的响应。例如,如果发现文件跨15个非相邻存储地址分布,则在重新编址之后,文件系统将其视为15个相邻存储地址位置。文件系统140然后可执行顺序访问以读取文件或向其写入,这比用以检索并组装每个碎片的15个离散事务快得多。本文所描述的实施例描述了存储设备160如何在不将文件碎片拷贝到可用的空闲储存器的情况下实现重新编址,并且替代地简单地将存储块重新编址到相邻的可寻址范围,使得设备驱动器150和文件系统140以更高效的传输方式操作。在其它实施例中,碎片整理应用120可以通过文件系统140和设备驱动器150简单地命令存储设备160应使用文件存储地址的供应列表来使得文件连续。如果该命令接收到成功响应,则文件系统知道其应使用(一个或多个)新的地址位置,而如果其接收到错误响应,则其可以在稍后重试重新编址。

[0037] 替换地,碎片整理应用120可以存在于非易失性存储设备160中,并且操作系统130可以命令非易失性存储设备160周期性地运行碎片整理应用120,并且非易失性存储设备160可以替代地在存储设备160本身上执行存储器的重新编址。在本示例中,由文件系统140向存储设备160提供信息,诸如文件列表和其被存储的碎片位置。在非易失性存储设备160

完成重新编址之后,其可以用描述文件内容的新位置的信息进行响应,并且将该改变上传到设备驱动器150和文件系统140。文件系统140然后将在读取和写入文件块时将新的地址用于在已重新编址位置处的文件碎片。

[0038] 在示例中,设备驱动器150可以包含碎片整理应用120,或者在设备驱动器外面的碎片整理应用120可以调用用以对非易失性存储设备160进行碎片整理或重新编址的例程。替换地,设备驱动器150可具有其自己的碎片整理应用120以开始重新编址操作以及通过将存储设备160通信地耦合到计算设备110的总线与被用来对存储位置重新编址的特殊协议命令通信。

[0039] 示例5—用于执行重新编址的方法

[0040] 图2是用于执行在非易失性存储设备160上的用于已破碎文件的存储器的重新编址的示例性方法200的流程图。

[0041] 在210处,向非易失性存储设备160发送用以对用于已破碎文件的存储器重新编址的命令。

[0042] 重新编址命令210的目标是将跨多个非连续寻址的存储块分布的文件(其基本上对于非易失性存储设备160而言表现为随机I/O访问模式)转换成较少(例如,一次)顺序访问。本文所描述的实施例实现对存储位置重新编址而不必将数据在物理上拷贝到新存储位置,将数据在物理上拷贝到新存储位置导致耗尽比重重新编址更多的功率,负面地影响存储器寿命,并且引入在对文件进行碎片整理为目的情况下将存储内容拷贝到操作系统并返回到存储器部分的明显更加冗长的I/O循环,其可能妨碍与最终用于想要运行的应用或正常操作系统行为相关联的任务。

[0043] 在220处,从非易失性存储设备160接收到用于已破碎文件的存储器已经被重新编址的响应。

[0044] 在230处,文件系统140将更新其内部记录保持,文件碎片在其中被寻址。在某些实施例中,文件系统140可以当在210处发送命令时更新其记录,并且如果其在220处未接收到成功响应,则回滚重新编址事务。在其它实施例中,文件系统140可等待直至其接收到基于在响应220中返回的新地址块实行或进行了相应重新编址改变的响应。例如,响应220可以包含用于在命令210中请求重新编址的块的新映射,并且用于块的最终协定的寻址在文件系统140在230处被更新时完成。

[0045] 在示例中,在文件系统140被更新之后,计算设备110可执行反映现在重新编址的存储器的操作。

[0046] 在示例中,计算设备可以使用包括相邻物理地址的现在重新编址的存储器来向非易失性存储设备160发送另一命令。例如,计算设备可以发送单个请求或多个请求的包以检索在相邻物理地址处的文件。因为文件位于相邻物理地址处,所以减少了用于计算设备的操作数目。可以更高效地利用非易失性存储设备160所使用的内部高速缓存机制,因为可以将重新编址之后的存储请求实现为对于数据的相邻顺序请求。相比于较小随机读和写而言,较大顺序读和写所固有的性能益处被当今存储设备的性能基准很好地表明,当今存储设备诸如是SD卡、eMMC设备、MMC以及SSD驱动。

[0047] 图3是用于执行在非易失性存储设备160上的用于已破碎文件的存储器的重新编址的示例性方法300的流程图。图3中所示的步骤对应于图2中所示的那些。在310处,接收用

以对已破碎文件的存储器进行重新编址的命令。

[0048] 在320处,向已破碎文件的存储器分配相邻物理存储器地址。也就是说,先前位于多个非相邻物理存储器地址处的文件碎片中的每个被重新编址到相邻物理存储器地址。

[0049] 在325处,非易失性存储设备160可返回处理重新编址改变的错误,并且系统将流到340,在该处不进行重新编址改变,并且中止重新编址。如果重新编址是成功的,则系统将流到330。如果非易失性存储设备160不能完成命令,则操作系统130可作为未重新编址的文件系统140的一部分而接收错误,如340所示。

[0050] 在330处,非易失性存储设备160可以用用于文件碎片的新地址位置对操作系统130(其包括设备驱动器150和文件系统140)进行响应。在某些实施例中,计算设备110可能不需要执行步骤330以对操作系统130进行响应,因为非易失性存储设备160简单地完成该命令。在其它实施例中,响应可能仅需要是块已经被重新编址的成功响应。

[0051] 在某些实施例中,可将重新编址逻辑包括为操作系统130的一部分,其跟踪所有块和可以被修改以进行重新编址碎片整理的可用块。在其它实施例中,操作系统可以请求非易失性存储设备160管理块,并且简单地要求对其知道严重破碎的文件进行重新编址,并且期待包含新块映射的响应。

[0052] 最后,在某些实施例中,重新编址将保持用于文件的原始起始块地址,并且重新编址将使得在起始地址之后寻址的所有后续存储块连续,使得其看起来是顺序访问;然而,后续块可能并不是实际上具有与可以计算为属于其它文件的地址相比较的唯一地址。稍后将在图4c中将其详细地描述为当两个文件所包含的块在外部观察者看起来是重叠块时结合了稀疏编址的重新编址解决方案。

[0053] 在相对于图1的示例性实施方式中,非易失性存储设备160可发送已破碎文件的存储器已被重新编址的响应330,但是响应被发送回去不是必要的。例如,非易失性存储设备160可仅接收用以对已破碎文件进行重新编址的命令,并且如果非易失性尺寸设备160正在正常地操作,则操作系统130、文件系统140、设备驱动器150或碎片整理应用120将假设其已成功地完成。

[0054] 例如,参考图1,用以对存储器重新编址的命令可来自文件系统140、设备驱动器150或非易失性存储设备160。在实施例中,可在计算设备110中存在另一操作系统部件,其向非易失性存储设备160提供命令。在另一实施例中,在操作系统130与非易失性存储器之间可存在单独部件,其向非易失性存储设备160提供命令。碎片整理应用120可存在于计算设备110、操作系统130、文件系统140和设备驱动器150中的一个或多个中。碎片整理从哪里发起的选择留给系统的设计者,其可基于各种供应商如何实现重新编址解决方案的质量和成本来选择要部署哪个应用模型。

[0055] 在相对于图1的示例性实施方式中,用以对存储器重新编址的命令被非易失性存储设备160接收到。然而,命令可能不指定需要对哪些已破碎文件重新编址。例如,该命令可以是到非易失性存储设备160的碎片整理请求的一部分。在这种情况下,非易失性存储设备160可基于文件的破碎程度来确定要重新编址的最可能候选文件,并选择该文件进行重新编址。然而,要重新编址的已破碎文件不需要是最破碎的文件。例如,非易失性存储设备160可基于文件被操作系统访问的频率、文件的物理存储器地址的位置或任何其它准则来确定最可能候选文件。可以由文件系统140或者按设备驱动器150或操作系统130或者甚至碎片

整理应用120所跟踪的来为非易失性存储设备160提供具有碎片的所有文件的列表。

[0056] 非易失性存储设备160可使用本文公开的任何方法来执行重新编址,但不限于那些方法。可执行对在非易失性存储设备上的用于已破碎文件的存储器重新编址的任何方法。

[0057] 示例6—示例性虚拟映射

[0058] 图4a和4b是示出了对物理存储器地址重新编址的示例的图。在本示例中,已破碎文件的文件碎片跨多个非相邻物理地址散布。例如,假设已破碎文件位于存储器地址1、3、4和7处。一旦存储设备接收到用以将已破碎文件重新编址到相邻物理存储器地址的命令,则存储设备确定将在哪些物理存储器地址处对文件进行重新编址。在本示例中,从存储器地址1处开始对文件重新编址,但是可替代地从任何物理存储器地址处开始对文件重新编址。

[0059] 在本示例中,将旧的存储器地址3重新编址到新的存储器地址2。然而,旧的存储器地址2可能包含其它数据。因此,存储器地址被交换,即旧的存储器地址3被重新编址到新的存储器地址2,并且旧的存储器地址2被重新编址到新的存储器地址3。针对已破碎文件的所有其余存储器地址重复此。最终结果是旧的存储器地址3、4和7被重新编址到新的存储器地址2、3和4,从而允许现在在相邻物理存储器地址处对已破碎文件的存储器进行寻址,并且旧的存储器地址2被重新编址到新的存储器地址7。在其中由文件系统140管理的簇或扇区尺寸与存储设备160中的块尺寸是紧密的1:1关系的情形中,本实施方式与图4a所示的非常类似。应认识到的是,对于在实际上并不拷贝数据的情况下对存储位置重新编址以便与实践已存的碎片整理解决方案相比以较少的数据拷贝和写入来执行碎片整理的基本原理而言,文件系统140所管理的扇区尺寸和簇尺寸不必要是1:1关系。

[0060] 然而,不一定需要交换物理存储器地址,而是替代地可以将其重新编址到未使用的存储器地址。例如,在图4b中,可将旧的存储器地址2重新编址到可用存储器地址100(例如,空的可用存储器地址)。已破碎文件的其它存储器地址然后能够被重新编址到相邻物理地址。

[0061] 图4c描述了保持用于文件的原始唯一起始块地址的重新编址机制的替换实施例。在本实施例中,重新编址使得在起始地址之后寻址的所有后续存储块连续,使得其在文件系统140或操作系统130看起来是顺序访问;然而,第二和后续块可能实际上并不具有与可以计算为属于其它文件的地址相比较的唯一地址(例如,第二和后续块可能具有可共享的物理存储器地址)。由于仅使用唯一起始地址和特定块长度来检索文件,并且由于其始终是顺序访问,所以在初始唯一块地址之后的内容流可以被明确地寻址。例如在图4c中,顶部的块集合示出了重新编址之前的状态;并且示出了文件“a.txt”,其在块1处开始,并且包含块3和5处的附加碎片达3个块的总长度(分别地如碎片a.txt₁、a.txt₂、a.txt₃所示)。该系统还可具有文件“b.txt”,其在块2处开始且具有作为碎片存储在块4和6中的内容(分别地示为b.txt₁、b.txt₂、b.txt₃)。用于文件“a.txt”的重新编址命令可以由文件系统140发送并命令对于块1、3、5处的文件的重新编址成为顺序的,即在块1处开始达到3个块的长度。此重新编址将使块1处的内容不改变,但是然后将块3重新编址为块2(仅当其在块1读取之后时)并将块5重新编址为块3(仅当其在块1和2的读取之后时)。因此,在重新编址之后,文件“a.txt”在被从外部源顺序地访问时被存储在块1、2和3中(如图4c的下半部上所示)。如果文件系统140将尝试针对单长度块而读取块2或3时,非易失性存储设备160将报告错误,因为其知道

系统必须仅通过读取块1和其后可选地读取块2和3作为单个顺序访问的一部分来检索文件。倘若使用在块1处开始的具有3的长度的命令顺序地寻址内容,内容被唯一地提供给文件系统40和存储驱动器150。文件系统140还命令对在块2处开始且包含块4和6的文件“b.txt”重新编址,假设其将使用如图4c中所示的类似机制变成块2、3、4。诸如文件系统140之类的外部观察者可计算到“a.txt”和“b.txt”两者的起始地址和长度共享块2和3中的内容,并且因此内容正在占用相同的存储块。然而,情况可能并非如此,因为非易失性存储设备160理解由于对“a.txt”和“b.txt”的访问始终是在唯一地址处开始的顺序访问,所以存储部分将在其接收到在块1处开始的3个块长的访问时输送仅被映射到文件“a.txt”的唯一内容,并且其将在其接收到在块2处开始的3个块长的请求时仅检索针对文件“b.txt”的内容。非易失性存储设备160将不向任何其它文件提供块地址1或2,并且在某些实施例中,文件系统140仅知道使用起始地址来访问文件,而不仔细检查重叠的文件和访问块。并且,在某些实施例中,文件系统140将不会接收到具有块3、4、5或6的起始地址的起始存储块,因为这些块实际上在被文件“a.txt”和“b.txt”使用。替换实施例可提供块3、4、5或6的起始地址;然而,提供给文件系统140的总地址块将不超过非易失性存储设备160的存储容量。此外,图4c示出了对空闲块7重新编址以提供3的可用起始块地址。假设文件系统140将存储新文件“c.txt”作为到地址3的1块写入,然后其将预期在其读取在地址2处开始的3块访问时接收到文件“b.txt”内容,并且预期将在其发布在地址3处开始的单块读取时获得“c.txt”内容。在某些实施例中,将仅对被以顺序方式读取的文件执行此类重新编址。在某些实施例中,文件系统140可以设定属性以不寻找以这种方式重新编址的这些种类的文件。

[0062] 应认识到的是,图4a-4c的示例全部示出了用于被文件系统140跟踪的文件碎片(即,簇)的离散存储单元与存储块(由非易失性存储设备160提供的离散存储单元,即,块或扇区)之间的1:1映射;然而,对于本文所描述的实施例而言,可以容易地显示文件碎片占用存储块的子部分以及存储在每个簇中的文件碎片跨多个可寻址存储块映射。也就是说,文件碎片(簇)和存储块(扇区)比可以是1:1、2:1、1:2、1:16、16:1等。简单地扩展了重新编址原理以确保在当需要移动部分页面时在重新编址处理期间进行最少的数据拷贝的情况下为存储提供唯一地址。并且,在重新编址之前和之后,由文件“a.txt”和“b.txt”消耗的存储空间将保持恒定,并且消耗非易失性存储设备160的6个块。

[0063] 在不移动非易失性存储设备160中的存储器的物位置的情况下执行物理存储器地址的重新编址(例如,如上文参考图4a和4b所描述)。图5a和5b是示出如何在不移动存储器的物理位置的情况下执行图4a和4b中所描述的重新编址操作的图。举先前的示例,旧的存储器地址1、3、4和7被重新编址到新的存储器地址1、2、3和4。然而,存储器设备上的已破碎文件的文件碎片的实际物理位置未被移动。参考本示例,图5a描绘了重新编址之前的存储器地址1、3、4和7。如图5a中所描绘的,非易失性存储设备160将对应于存储器地址1、3、4和7的文件碎片存储在非易失性存储设备160内的特定物理位置处,其在510处被以简化形式描绘为“LOC 1”至“LOC4”。图5b描绘了已执行存储器重新编址之后的存储器地址。如图5b中所描绘的,已执行重新编址,使得存储器地址现在是相邻的(地址1、2、3、4)。并且,如在图5b中所描绘的,即使已执行重新编址,非易失性存储设备160中的存储器的物理位置也未改变。因此,例如虽然旧的存储器地址3被重新编址到新的存储器地址2,但存储器的物理位置也未改变。

[0064] 在某些实施例中,执行此地址转换且支持重新映射的软件和/或硬件可以被存储在非易失性存储设备160内部。在其它实施例中,重新映射可以是跨文件系统140、存储驱动器150、以及非易失性存储设备160的分布式解决方案。例如,文件系统140可跟踪逻辑至物理块的映射,并且向非易失性存储设备160提交重新映射解决方案,其施加此改变。在其它实施例中,存储驱动器可以在其知道文件系统140已经将其映射到存储设备160中的存储块的地址之间执行转换,并且因此在文件系统140不知道重新映射的情况下为存储设备提供重新映射。

[0065] 示例7—示例性映射表

[0066] 图6a和6b是示出了图4a和4b中的已破碎文件的物理地址和LBA映射表的映射的示例。例如,LBA映射表可以被操作系统130用来向非易失性存储设备160上的物理地址分配逻辑地址。因为在先前的示例中物理地址已被重新编址,所以基于已重新编址的存储器来更新LBA映射表。因此,例如,在图6a中,LBA 0000指向物理地址1,LBA 0001指向物理地址3,LBA 0002指向物理地址4,并且LBA 0003指向物理地址7。在重新编址之后,如图6b中所示,LBA 0000指向物理地址1,LBA 0001指向物理地址2,LBA 0002指向物理地址3,并且LBA 0003指向物理地址4。

[0067] 如图6a和6b中所示,可更新LBA映射表以反映存储器的重新编址。然而,不一定需要更新LBA映射表。例如,在LBA映射表与存储设备之间可存在虚拟映射表。可用存储器的重新编址的新信息来更新虚拟映射表。当LBA映射表寻找地址时,已更新的虚拟映射表可指向已重新编址的物理地址,而LBA映射表不知道这样的重新编址已发生。在这种情况下,LBA映射表与包含用于已重新编址的物理存储器地址的信息的虚拟映射表通信。

[0068] 示例8—示例性计算环境

[0069] 图7描绘了其中可实现所描述创新的适当计算环境700的一般化示例。计算环境700并不意图暗示关于使用或功能性的范围的任何限制,因为可在各种各样的通用或专用计算系统中实现创新。例如,计算环境700可以是多种计算设备中的任何一个(例如,台式计算机、膝上型计算机、服务器计算机、平板计算机、媒体播放器、游戏系统、移动设备等)。

[0070] 参考图7,计算环境700包括一个或多个处理单元710、715和存储器720、725。在图7中,在虚线内包括此基本配置730。处理单元710、715执行计算机可执行指令。处理单元可以是通用中央处理单元(CPU)、专用集成电路(ASIC)中的处理器或任何其它类型的处理器。在多处理系统中,多处理单元执行计算机可执行指令以增加处理能力。例如,图7示出了中央处理单元710以及图形处理单元或协处理单元715。有形存储器720、725可以是易失性存储器(例如,寄存器、高速缓存、RAM)、非易失性存储器(例如,ROM、EEPROM、闪存等)或两者的某种组合,其可被(一个或多个)处理单元访问。存储器720、725以适合于由(一个或多个)处理单元执行的计算机可执行指令的形式存储实现本文所描述的一个或多个创新的软件780。

[0071] 计算系统可具有附加特征。例如,计算环境700包括储存器740、一个或多个输入设备750、一个或多个输出设备760以及一个或多个通信连接770。诸如总线、控制器或网络之类的互连机制(未示出)将计算环境700的部件互连。通常,操作系统软件(未示出)为在计算环境700中执行的其它软件提供操作环境,并且协调计算环境700的部件的活动。

[0072] 有形储存器740可以是可移除的或不可移除的,并且包括磁盘、磁带或盒、CD-ROM、

DVD或可以用来以非暂时方式存储信息且可以在计算环境700内被访问的任何其它介质。存储器740存储实现本文所描述的一个或多个创新的软件780的指令。

[0073] (一个或多个)输入设备750可以是触摸式输入设备,诸如键盘、鼠标、笔或轨迹球、语音输入设备、扫描设备或向计算环境700提供输入的另一设备。对于视频编码而言,(一个或多个)输入设备750可以是照相机、视频卡、TV调谐器卡或以模拟或数字形式接受视频输入的类似设备或者将视频样本读入到计算环境700中的CD-ROM或CD-RW。(一个或多个)输出设备760可以是显示器、打印机、扬声器、CD刻录机或从计算环境700提供输出的另一设备。

[0074] (一个或多个)通信连接770使得能够通过通信介质实现到另一计算实体的通信。通信介质在已调制数据信号中传送诸如计算机可执行指令、音频或视频输入或输出或者其它数据之类的信息。已调制数据信号是这样的信号,所述信号使其特性中的一个或多个被以这样的方式设定或改变以将信息编码在信号中。以示例而非限制的方式,通信介质可以使用电、光、RF或其它载体。

[0075] 可以在计算机可执行指令的一般上下文中描述该创新,诸如包括在程序模块中、在目标真实或虚拟处理器上的计算系统中执行的那些。一般地,程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、库、对象、类、部件、数据结构等。在各种实施例中可根据需要将程序模块的功能性被组合或在程序模块之间分离。可在本地或分布式计算系统内执行用于程序模块的计算机可执行指令。

[0076] 在本文中可互换地使用术语“系统”和“设备”。除非上下文清楚地指示,否则两个术语都不暗示对计算系统或计算设备的类型的任何限制。一般地,计算系统或计算设备可以是本地的或分布式的,并且可以包括具有实现本文所描述功能性的软件的专用硬件和/或通用硬件的任意组合。

[0077] 示例9—示例性移动设备

[0078] 图8是描绘包括一般地在802处所示的多种可选硬件和软件部件的示例性移动设备800的系统图。移动设备中的任何部件802可以与任何其它部件通信,但为了便于图示,并未示出所有连接。移动设备可以是多种计算设备中的任何一个(例如,蜂窝电话、智能电话、手持式计算机、个人数字助理(PDA)等),并且可以允许与一个或多个移动通信网络804的无线双向通信,诸如蜂窝、卫星或其它网络。

[0079] 所示移动设备800可以包括用于执行诸如信号编码、数据处理、输入/输出处理、功率控制和/或其它功能之类的任务的控制器或处理器810(例如,信号处理器、微处理器、ASIC或其它控制和处理逻辑电路)。操作系统812可以控制部件802的分配和使用并支持一个或多个应用814。应用814可以包括公共移动计算应用(例如,电子邮件应用、日历、联系人管理器、web浏览器、消息传送应用)或任何其它计算应用。用于访问应用商店的功能性813还可以用于获取和更新应用814。

[0080] 所示移动设备800可以包括存储器820。存储器820可以包括不可移除存储器822和/或可移除存储器824。不可移除存储器822可以包括RAM、ROM、闪存、硬盘或其它众所周知的存储器存储技术。可移除存储器824可以包括闪存或用户识别码模块(SIM)卡,其在GSM通信系统或其它众所周知的存储器存储技术中是众所周知的,诸如“智能卡”。存储器820可以用于存储用于运行操作系统812和应用814的数据和/或代码。示例性数据可以包括网页、文本、图像、声音文件、视频数据或将经由一个或多个有线或无线网络发送到一个或多个网络

服务器或其它设备和/或从其接收的其它数据集。存储器820可以用来存储诸如国际移动订户识别码(IMSI)之类的订户标识符和诸如国际移动设备标识符(IMEI)之类的设备标识符。可以将这样的标识符发射到网络服务器以识别用户和设备。

[0081] 移动设备800可以支持一个或多个输入设备830,诸如触摸屏832、麦克风834、照相机836、物理键盘838和/或轨迹球840以及一个或多个输出设备850,诸如扬声器852和显示器854。其它可能的输出设备(未示出)可以包括压电或其它触觉输出设备。某些设备可以服务于超过一个的输入/输出功能。例如,可以将触摸屏832和显示器854组合在单个输入/输出设备中。

[0082] 输入设备830可以包括自然用户接口(NUI)。NUI是使得用户能够以“自然”方式与设备相交互、从由诸如鼠标、键盘、遥控器等输入设备施加的人为约束解放出来的任何接口技术。NUI方法的示例包括依赖于语音识别、触摸和触针识别、在屏幕上的和邻近于屏幕的手势识别、空中手势、头和眼跟踪、语音和话音、视力、触摸、手势以及机器智能的那些。NUI的其它示例包括使用加速度计/陀螺仪的运动手势检测、人脸识别、3D显示器、头、眼、和视线跟踪、沉浸式加强现实和虚拟现实系统,其全部提供用于使用电场感测电极(EEG和相关方法)来感测大脑活动的更加自然的接口以及技术。因此,在一个特定示例中,操作系统812或应用814可以包括语音识别软件作为允许用户经由语音命令来操作设备800的语音用户接口的一部分。此外,设备800可以包括输入设备和软件,其允许经由用户的空间手势的用户交互,诸如检测并解释手势以向游戏应用提供输入。

[0083] 可以将无线调制解调器860耦合到天线(未示出),并且可以支持处理器810与外部设备之间的双向通信,如在本领域中很好地理解的。调制解调器860被一般地示出且其可以包括用于与移动通信网络804的通信的蜂窝调制解调器和/或其它基于无线电的调制解调器(例如,蓝牙864或Wi-Fi 862)。无线调制解调器860通常被配置用于与一个或多个蜂窝网络的通信,诸如用于在单个蜂窝网络内、蜂窝网络之间或者在移动设备与公共交换电话网(PSTN)之间的数据和语音通信的GSM网络。

[0084] 移动设备还可以包括至少一个输入/输出端口880、电源882、卫星导航系统接收机884,诸如全球定位系统(GPS)接收机、加速度计886和/或物理连接器890,其可以是USB端口、IEEE 1394(火线)端口和/或RS-232。所示部件802并不是要求的或包括一切的,因为可以删除任何部件,并且可以添加其它部件。

[0085] 示例10—示例性实施方式

[0086] 虽然为了便于呈现而按照特定的顺序次序描述了公开方法中的某些的实施例,应理解的是,这种描述方式涵盖重新布置,除非下面阐述的特定语言要求特定的排序。例如,顺序地描述的操作在某些情况下可被重新布置或同时地执行。而且,为了简单起见,附图可能并未示出可以结合其它方法而使用所公开方法的各种方式。

[0087] 可以将所公开方法的任一实现为存储在一个或多个计算机可读存储介质上并在计算设备(例如,任何可用的计算设备,包括智能电话或包括计算硬件的其它移动设备)上执行的计算机可执行指令。计算机可读存储介质是可以在计算环境内被访问的任何可用的有形介质(例如,非暂时计算机可读介质,诸如一个或多个光学介质磁盘、易失性存储器部件(诸如DRAM或SRAM)或非易失性存储器部件(诸如闪存或硬盘驱动))。以示例的方式并参考图7,计算机可读存储介质包括存储器720和725以及存储器740。以示例的方式并参考图

8,计算机可读存储介质包括存储器820、822以及824。术语计算机可读存储介质不包括诸如信号和载波之类的通信连接(例如,770、860、862和864)。

[0088] 可以将用于实现所公开的技术的任何计算机可执行指令以及在所公开实施例的实现期间创建和使用的任何数据存储在一个或多个计算机可读存储介质上。计算机可执行指令可以是例如专用软件应用或经由web浏览器或其它软件应用(诸如远程计算应用)访问或下载的软件应用的一部分。这样软件可以例如在单个本地计算机(例如,任何适当的商业上可得到的计算机)上或者使用一个或多个网络计算机在网络环境中(例如,经由因特网、广域网、局域网、客户端-服务器网络(诸如云计算网络)或其它这样的网络)执行。

[0089] 为了清楚起见,仅描述了基于软件的实施方式的特定的所选方面。省略了在本领域中众所周知的其它细节。例如,应理解的是,所公开的技术不限于任何特定计算机语言或程序。例如,可以由C++、Java、Perl、JavaScript、Adobe Flash或任何其它适当编程语言编写的软件来实现所公开的技术。同样地,所公开的技术不限于任何特定计算机或硬件类型。适当计算机和硬件的特定细节是众所周知的,并且在本公开中不需要详细地阐述。

[0090] 还应很好地理解的是,本文所描述的任何功能性可以至少部分地由一个或多个硬件逻辑部件而不是软件执行。例如,并且在没有限制的情况下,可以使用的说明性类型的硬件逻辑部件包括现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SOC)、复杂可编程逻辑设备(CPLD)等。

[0091] 此外,可以通过适当的通信手段来上传、下载或远程地访问任何基于软件的实施例(包括例如用于促使计算机执行任何所公开的方法的计算机可执行指令)。这样的适当的通信手段包括例如因特网、万维网、内联网、软件应用、电缆(包括光纤电缆)、磁通信、电磁通信(包括RF、微波以及红外通信)、电子通信或其它这样的通信手段。

[0092] 不应将公开的方法、装置以及系统理解为以任何方式进行限制。替代地,本公开针对各种公开的实施例的所有新颖和非显而易见的特征和方面(单独地和以相互的各种组合和子组合的方式)。所公开的方法、装置以及系统不限于任何特定方面或特征或其组合,所公开的实施例也不要求存在任何一个或多个特定优点或者解决任何一个或多个特定问题。

[0093] 替换方案

[0094] 可以将来自任何示例的技术与在其它示例中的任何一个或多个中描述的技术组合。鉴于所公开的技术的原理可被应用到的许多可能实施例,应认识到的是所示示例是所公开的技术的示例,并且不应理解为对所公开的技术的范围的限制。相反地,所公开的技术的范围包括以下权利要求涵盖的内容。因此我们要求进入这些权利要求的范围内的所有作为我们的发明。

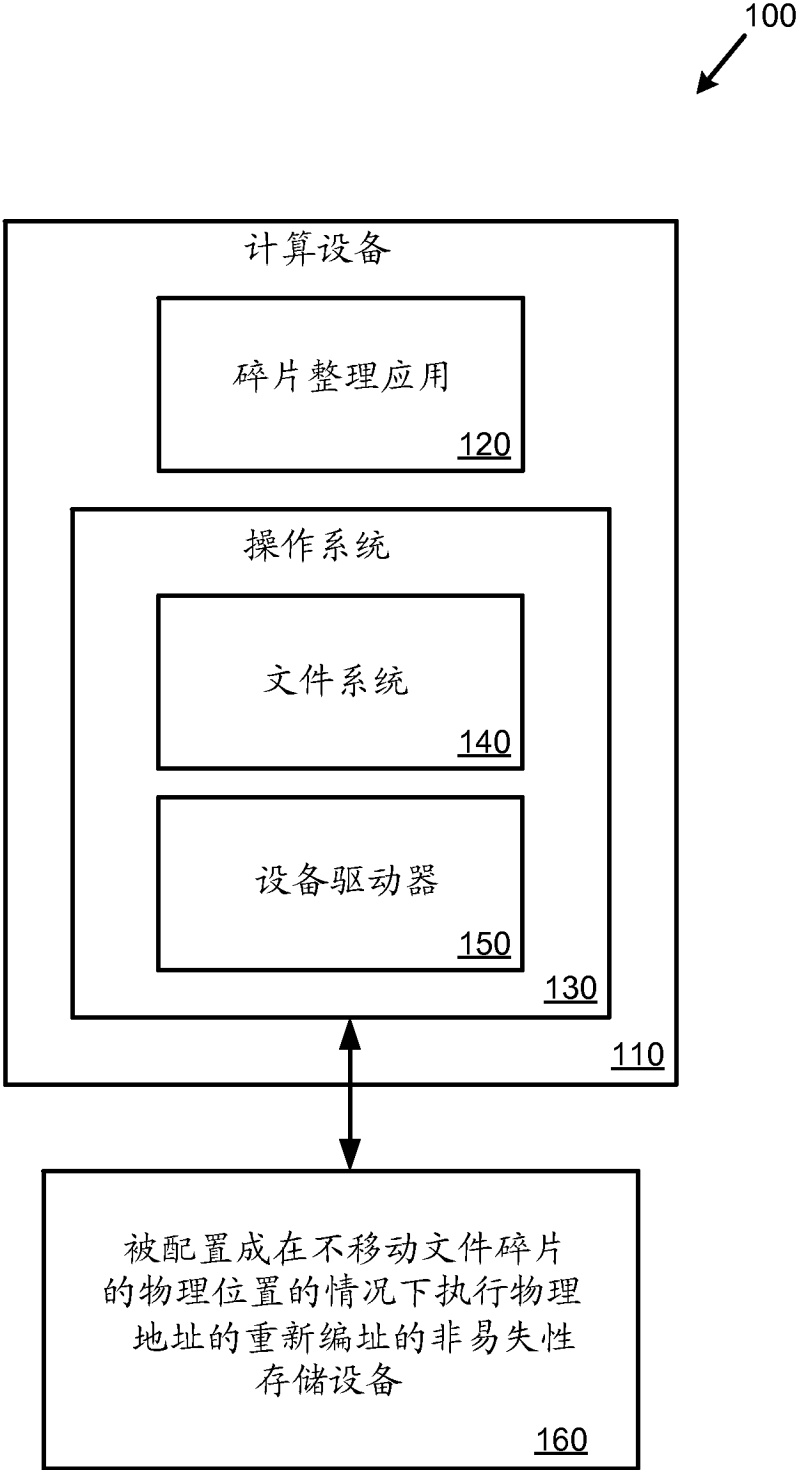


图 1

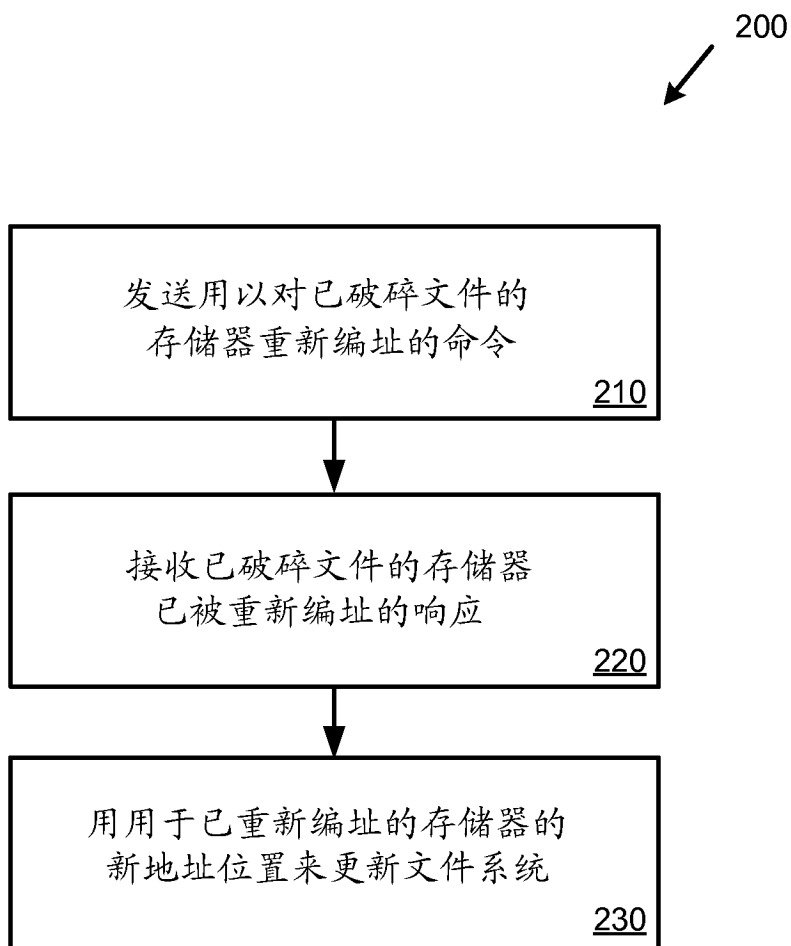


图 2

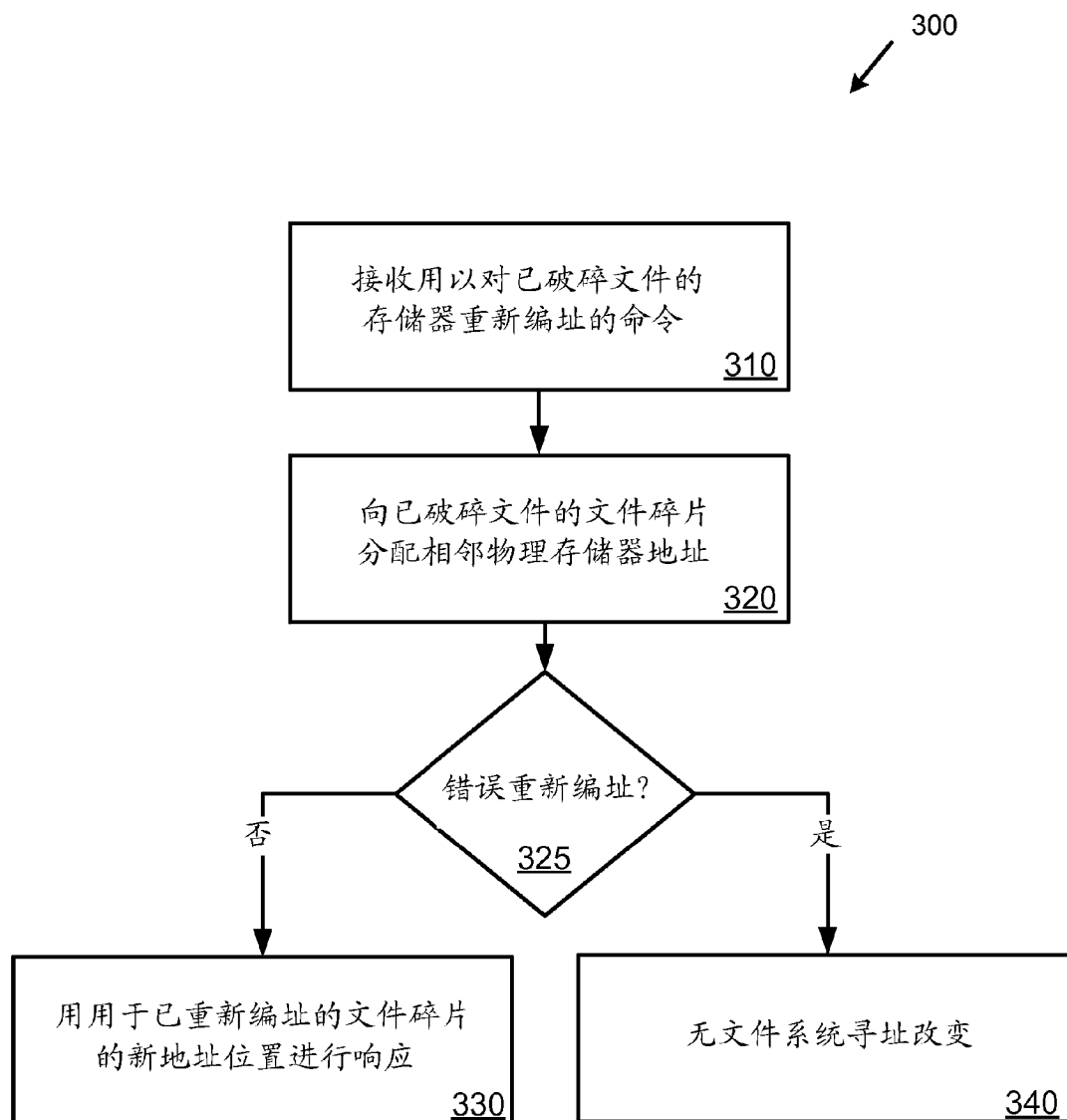


图 3

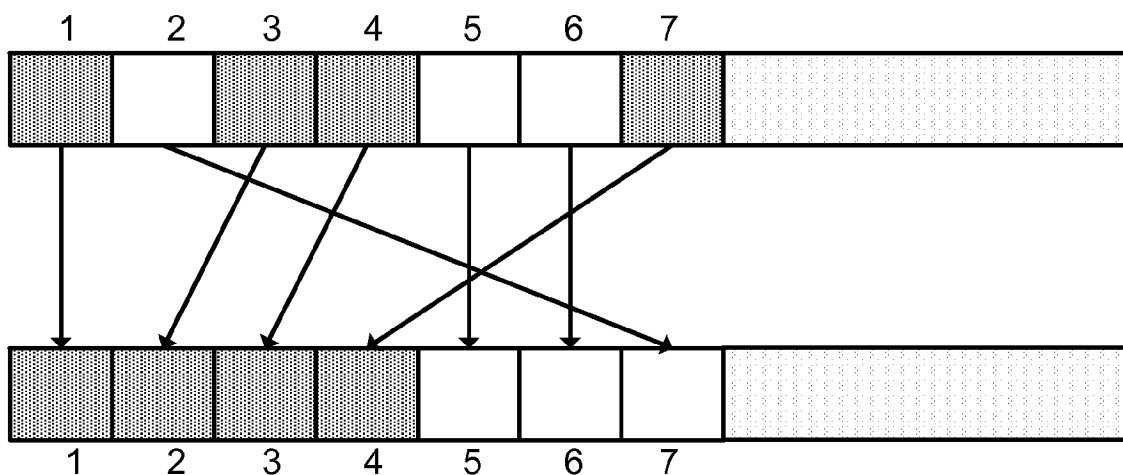


图 4a

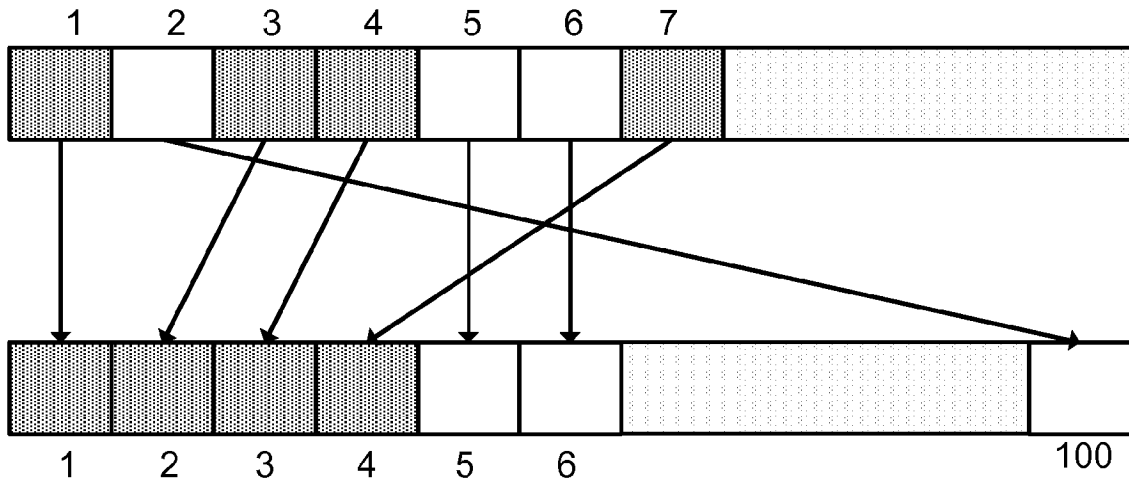


图 4b

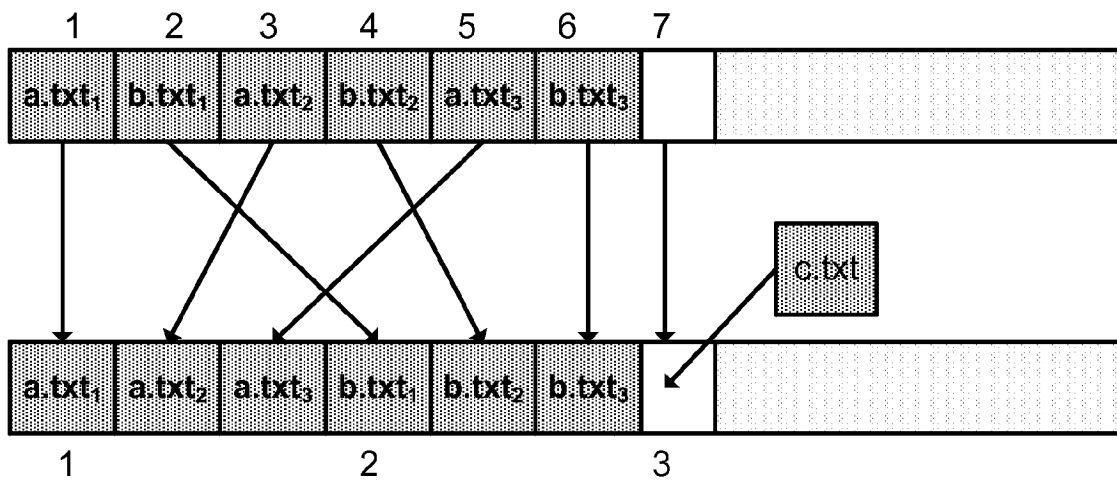


图 4c

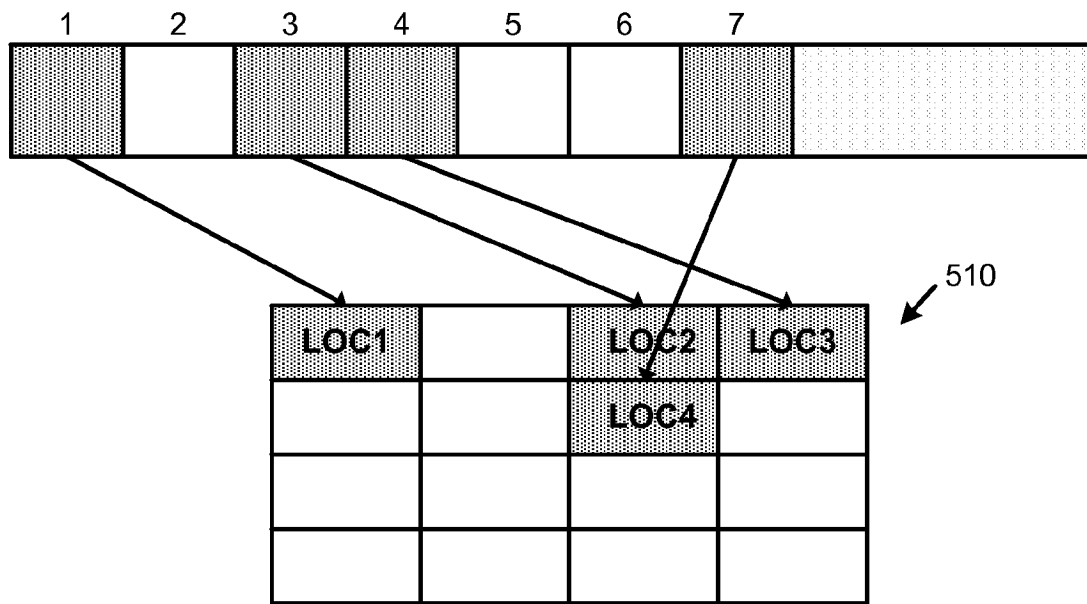


图 5a

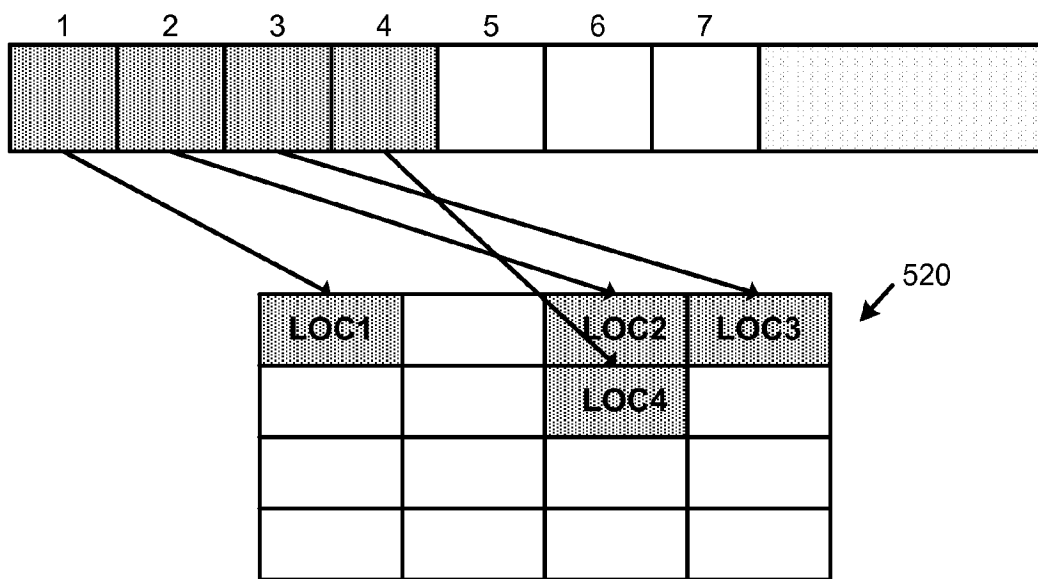


图 5b

LBA	物理地址
0000	1
0001	3
0002	4
0003	7

图 6a

LBA	物理地址
0000	1
0001	2
0002	3
0003	4

图 6b

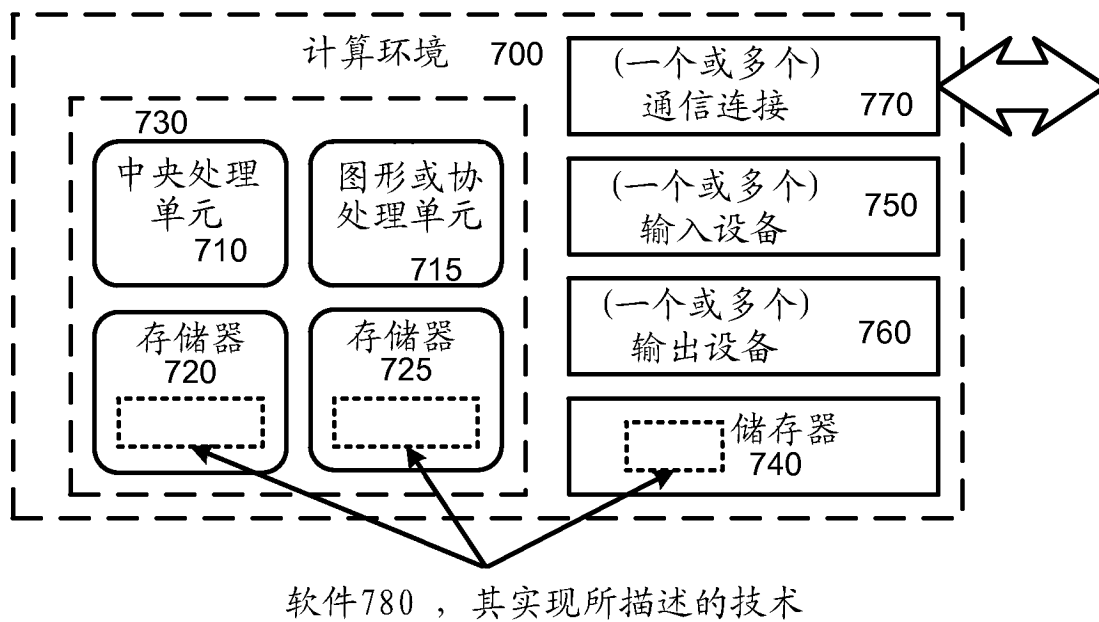


图 7

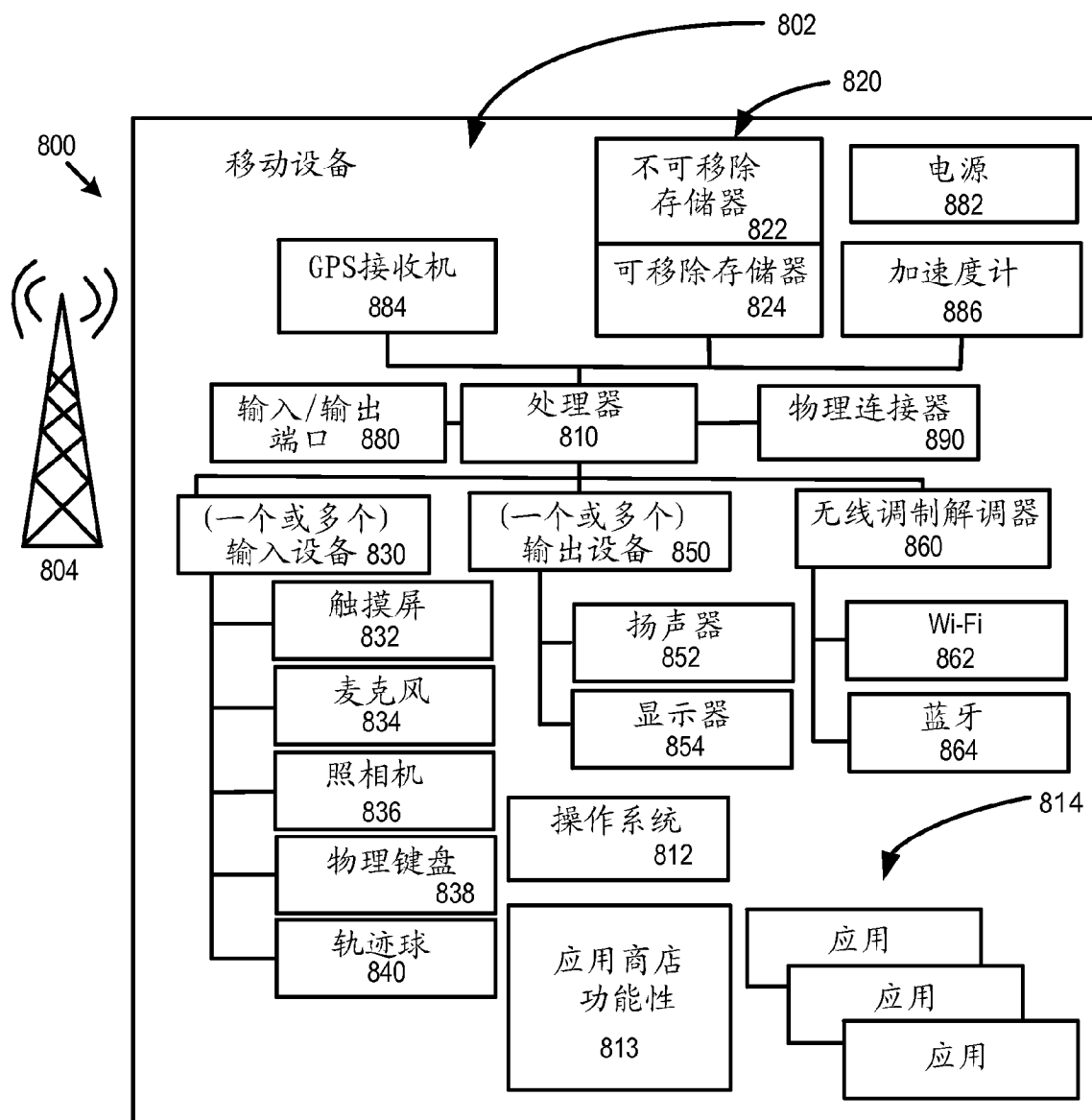


图 8