



(12) 发明专利申请

(10) 申请公布号 CN 102884516 A

(43) 申请公布日 2013. 01. 16

(21) 申请号 201080066763. 6

G06F 3/06 (2006. 01)

(22) 申请日 2010. 05. 13

(85) PCT申请进入国家阶段日
2012. 11. 13

(86) PCT申请的申请数据
PCT/US2010/034794 2010. 05. 13

(87) PCT申请的公布数据
W02011/142762 EN 2011. 11. 17

(71) 申请人 惠普发展公司, 有限合伙企业
地址 美国德克萨斯州

(72) 发明人 K. K. 桑帕思库马

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

代理人 段俊峰 卢江

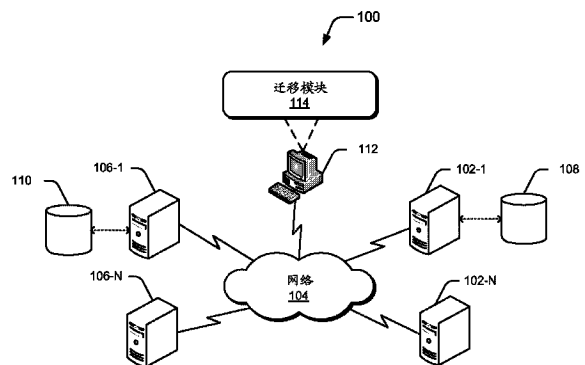
(51) Int. Cl.
G06F 15/17 (2006. 01)
G06F 15/76 (2006. 01)

权利要求书 2 页 说明书 11 页 附图 6 页

(54) 发明名称
文件系统迁移

(57) 摘要

描述了一种用以执行文件系统迁移的方法。该方法包括使源块设备关联至目的地块设备, 其中, 所述源块设备和目的地块设备分别包括多个源块和目的地块。用于来自所述多个源块的源块的至少一个命令至少基于该关联被引导到源块设备或目的地块设备。此外, 部分地基于所述至少一个所引导的命令来更新来自所述多个目的地块的目的地块。



1. 一种用以执行文件系统迁移的方法,该方法包括:
使源块设备关联至目的地块设备,其中,所述源块设备和目的地块设备分别包括多个源块和目的地块;
至少基于所述关联将用于来自所述多个源块的源块的至少一个命令引导到所述源块设备或目的地块设备;以及
部分地基于所述至少一个所引导命令来更新来自所述多个目的地块中的目的地块。
2. 如权利要求 1 所述的方法,其中,所述至少一个命令是在文件系统迁移期间生成的。
3. 如权利要求 1 所述的方法,其中,引导所述至少一个命令还包括随后在所述至少一个命令是写请求时将所述至少一个命令引导到与源块相关联的目的地块。
4. 如权利要求 1 所述的方法,其中,所述引导还包括:
针对所述至少一个命令是读请求,随后将所述至少一个命令引导到源块,如果对应的目的地块未被更新的话;以及
针对所述至少一个命令是读请求,随后将所述至少一个命令引导到来自所述多个目的地块的对应于源块的另一目的地块,如果所述另一目的地块已被更新的话。
5. 如权利要求 1 所述的方法,还包括将与未被更新的所述多个目的地块相关联的所述多个源块拷贝到目的地块设备。
6. 一种计算设备,包括:
处理器;
存储器,其被耦合至所述处理器,其中,所述存储器包括;
关联模块,其被配置成使源块设备与目的地块设备相关联;以及
迁移模块,其被配置成:
确定与所述源块设备上的至少一个块相关联的所述目的地块设备上的至少一个块是否在所述目的地块设备上被更新;以及
基于所述确定使所述源块设备上的文件系统迁移至所述目的地块设备。
7. 如权利要求 6 所述的计算设备,其中,所述迁移模块还被配置成修改文件系统更新位图以跟踪对目的地块设备上的所述至少一个块的更新。
8. 如权利要求 7 所述的计算设备,其中,所述迁移模块还被配置成当目的地块设备上的所述至少一个块被更新时修改文件系统更新位图。
9. 如权利要求 6 所述的计算设备,其中,所述迁移模块还被配置成计算指示在目的地块设备处未被更新的源块设备上的所述至少一个块的拷贝位图。
10. 如权利要求 9 所述的计算设备,其中,所述迁移模块还被配置成基于拷贝位图来迁移所述文件系统。
11. 一种具有在被执行时执行动作的一组计算机可读指令的计算机可读介质,所述动作包括:
响应于用于源块设备上的至少一个块的重新定位的至少一个命令,基于源块设备与目的地块设备之间的关联将来自源块设备的块重新定位至目的地块设备;以及
响应于所述重新定位来更新与源块设备相关联的文件系统更新位图;
基于更新将来自源块设备的至少一个其他块拷贝至目的地块设备,其中,所述其他块未被重新定位。

12. 如权利要求 11 所述的计算机可读介质,其中,所述拷贝还包括基于文件系统更新位图和目的地分配位图来计算拷贝位图。

13. 如权利要求 11 所述的计算机可读介质,其中,当源块设备的块中没有被更新时,所述拷贝是基于源分配位图的。

14. 如权利要求 11 所述的计算机可读介质,其中,所述重新定位包括将源块设备的所述至少一个块映射到目的地块设备上的至少一个块。

15. 如权利要求 14 所述的计算机可读介质,其中,所述重新定位还包括当至少一个块在目的地块设备处未被更新时基于所述映射将命令引导至所述至少一个源块设备以便从源块设备读取所述至少一个块。

文件系统迁移

背景技术

[0001] 一般地,文件系统的迁移涉及将位于源设备的诸如盘分区的源块设备上的文件系统拷贝到目的地设备的目的地块设备并复制该文件系统。块设备还可以包括例如逻辑盘或逻辑卷。

[0002] 在文件系统迁移期间,可以在目的地块设备处修改文件系统。可以出于各种目的进行对文件系统的修改,例如以保证存储合并、以使得空间可用于未来存储以及以实现有效的盘空间利用。可以存在某些修改,诸如缩小文件系统的大小,其在目的地块设备处可能不是可实现的。文件系统的此类修改可以在源块设备与目的地块设备之间的中间存储设备处或在源块设备处执行。

[0003] 一般地,在文件系统迁移期间即使是临时的也不存在可用的中间存储设备。另外,可以将源块设备置于静默状态,即其中对源块设备上的文件系统的访问可用于只读目的且不能在源块设备上对文件系统进行修改的状态。源块设备被置于此类状态以参考源块设备来保持目的地块设备处的文件系统的一致性和完整性。

附图说明

[0004] 参考附图来描述详细描述。在图中,附图标记的(一个或多个)最左数字识别其中附图标记第一次出现在其中的图。遍及各图使用相同的数字来指代相同的特征和部件。

[0005] 图 1 说明根据本发明的实施方式的具有用于执行文件系统迁移的一个或多个设备的示例性网络环境。

[0006] 图 2 说明根据本发明的实施方式的用于执行文件系统迁移的示例性设备。

[0007] 图 3 说明根据本发明的实施方式的示例性文件系统更新位图。

[0008] 图 4 说明根据本发明的实施方式的示例性分配位图。

[0009] 图 5 说明根据本主题的实施方式的用于处理文件系统命令的示例性方法。

[0010] 图 6 说明根据本主题的实施方式的用于执行文件系统块的迁移的示例性方法。

具体实施方式

[0011] 在本文中描述了用于执行文件系统迁移的方法和设备。可以使文件系统从源设备处的源块设备迁移至目的地设备处的目的地块设备。块设备上的文件系统可以采取称为文件系统块的分立数据块的形式。文件系统在迁移期间涉及文件系统块从源块设备到目的地块设备的转移。

[0012] 从源块设备到目的地块设备的文件系统迁移的过程可以包括文件系统的修改,例如重新确定文件系统的大小。文件系统重新确定大小涉及调整文件系统在块设备上利用的盘空间。在此类修改期间,可以根据目的地块设备上可用的盘空间放大或缩小文件系统的大小。

[0013] 文件系统的某些修改、诸如放大文件系统的大小一般被实现为目的地块设备处的迁移后步骤。在这种情况下,在文件系统迁移到目的地块设备上之后,可以运行文件系统工

具以将文件系统扩展至目的地块设备的新大小。此工具保证文件系统的大小被放大以扩展并占用目的地块设备上的附加空间。

[0014] 可以以两个方式来实现诸如缩小文件系统的大小的修改,所述方式即具有非重新定位的文件系统修改和具有重新定位的文件系统修改。如本领域中已知的,重新定位涉及使文件系统中的数据块移动,例如以在块设备上产生邻接的自由空间。

[0015] 在具有非重新定位的文件系统修改中,在目的地块设备处文件系统块未被重新定位。另外,在缩小文件系统的大小期间,文件系统的文件系统块中的数据保持不受影响。具有有效数据的文件系统最后文件系统块与文件系统的结尾之间的任何邻接自由空间确定能够实现的文件系统的大小缩小的程度。此外,在本情况下,未实现文件系统的碎片整理,并且还限制了大小缩小的程度。

[0016] 另一方面,在具有重新定位的文件系统修改中,在给定块设备上重新定位文件系统块。在数据的重新定位期间,文件系统块中的元数据和对应的数据被到处移动以组合在块设备上的各部分中分散的自由空间。使得文件系统块中的数据在块设备的一个部分中邻接,并且在块设备上使得邻接自由空间可用以未来存储。此外,通常朝着块设备上的存储空间的结尾使得邻接自由空间可用。用这种方法,能够实现文件系统的有效的大小缩小。

[0017] 然而,在重新定位文件系统块的情况下在目的地块设备处缩小文件系统的大小在文件系统块的迁移期间难以处理。当对应目的地块设备处的空间小于源块设备上的文件系统所拥有的空间时,困难可能是明显的。目的地块设备处的空间约束可以归因于各种原因,例如以保证存储合并、以使得空间可用于未来存储以及以实现有效的块设备空间利用。在这种情况下,文件系统的修改在目的地块设备处可能是不可实现的。

[0018] 此外,在文件系统迁移期间,中间存储设备不可用于缓冲数据并执行任何修改或变更以作为文件系统迁移的一部分。

[0019] 为此,描述了用以执行文件系统迁移的系统和方法。在一个实施方式中,通过使源块设备和目的地块设备相关联来创建映射对。该关联是通过使源块设备上的块映射到目的地块设备上的对应块而实现的。

[0020] 在一个实施方式中,该关联用来捕捉用于块的更新并作为文件系统迁移过程的一部分使其改向(redirect)至目的地块设备。例如,监视诸如读或写请求的命令。如果接收到意图用于源块设备的块中的一个的写请求,则该写请求随后被引导到目的地块设备上的对应块。响应于写请求,更新目的地块设备上的块。基于源块设备和目的地块设备之间的关联、即基于映射对来确定目的地块设备上的块。因此,使用于文件系统的更新改向至目的地块设备。

[0021] 另一方面,如果接收到意图用于源块设备的块中的一个的读请求,则进行确定以检查要被读取的块是否已被更新。如果源块设备上的该块尚未被更新,则随后使读请求引导到源块设备的块。然而,如果源块设备上的块在目的地块设备上已被更新,即被拷贝并重新定位,则随后该读请求被引导到目的地块设备上的已更新块。因此,使对于在目的地块设备处已被更新的源块设备上的块的读请求改向至目的地块设备上的对应块。

[0022] 在一个实施方式中,可以使用文件系统更新位图来跟踪对目的地块设备上的块的更新。在一个实施方式中,可以用阵列来表示文件系统更新位图,其中,每个位可以对应于目的地块设备上的块。例如,文件系统更新位图中的位在设定时(即 1)将指示目的地块

设备上的对应块已被更新。相反,某个位的清零值(即当为 0 时)将指示对应的块未被更新。

[0023] 如前所述,形成文件系统迁移的一部分的更新和修改被引导到目的地块设备。在这种情况下,在目的地块设备处存在除了具有数据并因此被分配的未更新块以外的所有块。因此,为了完成迁移过程,具有有效数据但未被更新的那些块被从源块设备转移至目的地块设备。

[0024] 为此,基于文件系统更新位图来确定目的地块设备上的块的更新状态。在一个实施方式中,基于拷贝位图从源块设备拷贝块。此外,可以针对被分配但未被更新的源块设备处的块(即对于其而言文件系统更新位图的对应的块被清零的块)来计算拷贝位图。在一个实施方式中,拷贝位图基于文件系统更新位图和目的地分配位图。在所述实施方式中,针对对应于文件系统更新位图中的清零位的那些块来计算拷贝位图。

[0025] 可以理解的是当文件系统的块被分成若干块组时,可以以与如上所述类似的方式来实现文件系统迁移。在这种情况下,基于文件系统更新位图中的更新状态字段来确定块组的更新状态。如果用于块组的更新状态被清零,则其指示该块组中的块没有一个被更新。在这种情况下,基于块组的源分配位图从源块设备拷贝块。

[0026] 另一方面,用于块组的更新状态在被设定时指示块组中的块中的至少一个被更新。现在将拷贝源块设备上未被更新但被分配在源块组中的块。为了识别此类块,针对该块组来计算拷贝位图。在一个实施方式中,拷贝位图基于块组更新位图和目的地块组位图。

[0027] 本文所述的方法和设备在动态重新定位目的地块设备上的文件系统块的情况下执行文件系统迁移,而不影响源块设备上的文件系统。此外,此类文件系统迁移实现数据的有效合并、创建用于未来存储的空间并实现自由空间的高效利用。

[0028] 应关于图 1 至图 6 来详细地解释用以执行文件系统迁移的方式。虽然可以在任何数目的不同计算系统、环境和 / 或配置中实现系统和方法的各方面,但在以下(一个或多个)示例性系统架构的上下文中描述实施例。

[0029] 图 1 说明根据本发明的实施方式的用于执行文件系统迁移的示例性网络环境 100。网络环境 100 可以包括一个或多个源设备 102-1、102-2、...102-N (共同称为源设备 102)。源设备 102 被连接至网络 104。网络环境 100 还可以包括一个或多个目的地设备 106-1、106-2、...106-N (共同称为目的地设备 106)。

[0030] 源设备 102 和目的地设备 106 可以是多种设备,诸如服务器、主机计算机、个人计算机、膝上型计算机、个人数字助理(PDA)等。例如,在一个实施方式中,网络环境 100 可以是公司网络,包括几千个办公室个人计算机、膝上型计算机、各种服务器,诸如刀片服务器以及通过网络 104 连接的其他计算设备。在另一实施方式中,网络环境 100 可以是具有通过网络 104 连接的有限数目的个人计算机和膝上型计算机的家庭网络。

[0031] 目的地设备 106 可以通过网络 104 与源设备 102 通信。网络 104 可以是无线网络、有线网络或其组合。可以将网络 104 实现为不同类型的网络中的一个,诸如内联网、局域网(LAN)、广域网(WAN)、互联网等。网络 104 可以是专用网络或共享网络,其表示使用多种协议的不同类型网络的关联以相互通信,所述多种协议例如超文本传输协议(HTTP)、传输控制协议 / 网际协议(TCP/IP)、无线应用协议(WAP)等。

[0032] 此外,源设备 102 中的每一个与源块设备 108 相关联。源块设备 108 可以具有与之相关联的对应分配位图,在下文中称为源分配位图。源分配位图可以指示源块设备 108

上的块的分配状态。

[0033] 此外,可以使目的地设备 106 中的每一个与目的地块设备 110 相关联。目的地块设备 110 可以具有与之相关联的对应分配位图,称为目的地分配位图。目的地分配位图可以指示目的地块设备 110 上的块的分配状态。

[0034] 源块设备 108 和目的地块设备 110 可以是例如物理盘分区或逻辑盘分区,诸如逻辑卷。在一个实施方式中,块设备 108、110 可以分别在源设备 102 和目的地设备 106 内部。在另一实施方式中,源块设备 108 和目的地块设备 110 可以是驻留于在网络环境 100 中示出的设备中的任何一个中的同一物理块设备的两个分离的逻辑分区。

[0035] 在一个实施方式中,源块设备 108 存储将被迁移至目的地块设备 110 的文件系统。可以以多个数据块的形式将文件系统存储在源块设备 108 上。此外,可以将所述多个块聚集在一起以形成一个或多个块组。此外,目的地块设备 110 可以具有与源块设备 108 类似的存储配置。

[0036] 根据实施方式,网络环境 100 还包括一个或多个计算设备 112,例如管理员计算设备,以执行文件系统迁移。可以理解的是虽然计算设备 112 被描绘为具有与源设备 102 和目的地设备 106 不同的配置,但其可以具有与源设备 102 或目的地设备 106 之一相同或类似的配置。可以将计算设备 112 实现为服务器、主机计算机、个人计算机、膝上型计算机、个人数字助理(PDA)等。

[0037] 在一个实施方式中,计算设备 112 包括迁移模块 114。在文件系统迁移期间,可以由迁移模块 114 将位于源块设备 108 上的文件系统拷贝至所识别的目的地块设备 110。可以理解的是虽然迁移模块 114 被示为驻留于计算设备 112 上,但迁移模块 114 可以驻留于源设备 102 或目的地设备 106 上。

[0038] 在一个实施方式中,通过使源块设备 108 和目的地块设备 110 相关联、从而创建映射对来执行文件系统迁移。迁移模块 114 发布例如读或写请求的命令作为文件系统迁移过程的一部分。通过源块设备 108 与目的地块设备 110 之间的关联,捕捉并监视由迁移模块 114 发布的命令。

[0039] 此外,该命令随后基于监视而被引导到相关块设备。例如,意图用于源块设备 108 上的块中的一个的写请求被改向至目的地块设备 110 上的对应块且目的地块设备 110 上的块被更新。目的地块设备 110 上的块的更新包括例如从源块设备 108 拷贝块和将所拷贝块重新定位到目的地块设备 110 上。此外,执行目的地块设备 110 上的块的此类更新作为文件系统迁移的一部分。在一个实施方式中,基于源块设备 108 和目的地块设备 110 的关联来确定目的地块设备 110 上的对应块。

[0040] 类似地,当接收到意图用于源块设备 108 的块中的一个的读请求时,在计算设备 112 处进行确定。该确定查明要被读取的块在目的地块设备 110 上是否已被更新。如果该块尚未被更新,则读请求被引导到源块设备 108 上的块。然而,如果源块设备 108 上的块已被更新,则使读请求改向至目的地设备 106 上的已更新块。

[0041] 如上所述,用于源块设备 108 上的已更新块的读请求随后被引导到目的地块设备 110 上的关联块。为了处理由迁移模块 114 生成的命令,在一个实施方式中,可以使用文件系统更新位图来跟踪对目的地块设备 110 上的块的更新。在一个实施方式中,文件系统更新位图中的位可以对应于目的地块设备 110 上的块。例如,文件系统更新位图的块在被设

定时(即 1)将指示目的地块设备 110 上的对应块已被更新。相反,用于某个块的清零值(即 0)将指示对应的块未被更新。

[0042] 为了完成文件系统的迁移,迁移模块 114 然后继续基于文件系统更新位图来识别具有有效数据但未被更新的源块设备 108 上的那些块。在一个实施方式中,可以基于文件系统更新位图的更新状态从源块设备 108 识别剩余要拷贝的块。

[0043] 在另一实施方式中,将源块设备 108 和目的地块设备 110 的块结构化为块组。在这种情况下,文件系统更新位图可以包括更新状态字段。更新状态字段在被设定时指示块组内的至少一个块在目的地块设备 110 上已被更新。例如,块组内的一百个块之中的一个在被更新时将设置用于该块组的更新状态字段。类似地,用于块组的更新状态字段在清零时指示该块组内的块中没有一个是已被更新。文件系统更新位图还可以包括块组更新位图,其对应于目的地块设备 110 上的块组。在一个实施方式中,块组更新位图指示目的地块设备 110 上的块组中的各个块的更新状态。

[0044] 结合图 2 来进一步详细地描述用以实现文件系统迁移的方式。图 2 说明按照本主题的一个实施方式的用于执行文件系统迁移的计算设备 112 的示例性部件。

[0045] 在一个实施方式中,计算设备 112 包括被耦合至存储器 204 的(一个或多个)处理器 202。可以将该(一个或多个)处理器 202 实现为一个或多个微处理器、微型计算机、微控制器、数字信号处理器、中央处理单元、状态机、逻辑电路和 / 或基于操作指令来操纵信号的任何设备。除其他能力之外,该(一个或多个)处理器 202 被配置成获取并执行存储在存储器 204 中的计算机可读指令。

[0046] 存储器 204 可以包括在本领域中已知的计算机可读介质,包括例如易失性存储器,诸如静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)等和 / 或非易失性存储器,诸如可擦编程只读存储器(EPROM)、闪存存储器等。在一个实施方式中,存储器 204 包括(一个或多个)模块 206 和数据 208。

[0047] 该(一个或多个)模块 206 包括迁移模块 114、关联模块 210、命令驱动器 212 以及(一个或多个)其他模块 214。(一个或多个)其他模块 214 可以包括补充由计算设备 112 执行的应用或功能的程序或编码指令。

[0048] 数据 208 包括文件系统更新位图 216、用于文件系统的每个块组的拷贝位图 218、源分配位图 220、目的地分配位图 222 以及其他数据 224。其他数据 224 还可以包括被(一个或多个)模块 206 中的一个或多个模块使用或作为其执行结果而生成的数据。

[0049] 可以分别用源分配位图 220 和目的地分配位图 222 来表示源块设备 108 和目的地块设备 110 内的数据。分配位图 220 和 222 两者分别指示源块设备 108 和目的地块设备 110 的块的分配状态。

[0050] 计算设备 112 还可以包括(一个或多个)接口 226,例如以促进与源块设备 108 和目的地块设备 110 的通信。该(一个或多个)接口 226 可以包括多种软件和硬件接口,例如用于诸如键盘、鼠标、外部存储器以及打印机的(一个或多个)外围设备的接口。此外,该(一个或多个)接口 226 可以使得计算设备 112 能够与其他计算设备通信,诸如源设备 102、目的地设备 106 以及外部数据库。该(一个或多个)接口 226 还可以促进多种网络和协议类型内的多个通信,包括例如 LAN、电缆等的有线网络和诸如 WLAN、蜂窝或卫星的无线网络。为了该目的,该(一个或多个)接口 226 可以包括用于将若干计算设备相互连接或连接至其他

服务器计算机的一个或多个端口。

[0051] 在操作中,计算设备 112 执行从源块设备 108 至目的地块设备 110 的文件系统迁移。在一个实施方式中,文件系统迁移包括块到目的地块设备 110 上的重新定位以作为文件系统迁移期间的文件系统的动态修改的一部分。

[0052] 为此,关联模块 210 将源块设备 108 关联或映射至目的地块设备 110 以创建映射对。在一个实施方式中,该映射对可以包括描述源块设备 108 和目的地块设备 110 的映射的信息。可以将此信息存储在其他数据 224 中。在一个实施方式中,所述映射信息描述源块设备的块和 / 或块组与目的地块设备 110 的对应块和 / 或块组之间的映射。

[0053] 在一个实施方式中,关联模块 210 在创建映射对时创建文件系统更新位图 216。文件系统更新位图 216 跟踪目的地块设备 110 中的各种块的更新状态。例如,文件系统更新位图 216 中的位对应于目的地块设备 110 上的块。此外,如前所述,在块被布置成块组的情况下,文件系统更新位图 216 还可以包括关于目的地块设备 110 上的文件系统中的各种块组的更新状态的信息。例如,文件系统更新位图 216 中的更新状态字段可以跟踪目的地块设备 110 中的块组的更新状态。类似地,一个或多个块组更新位图可以跟踪对应块组中的块的更新状态。在图 2 中文件系统更新位图 216 已被示为在计算设备 112 上。然而,在不脱离本主题的范围的情况下,文件系统更新位图 216 还可以存在于源设备 102 中。在这种情况下,迁移模块 114 将文件系统更新位图 216 从源设备 102 导入到计算设备 112 或通过(一个或多个)接口 226 来访问源设备 102 上的文件系统更新位图 216。

[0054] 在文件系统迁移过程期间,迁移模块 114 可以向块、例如向具有数据的块发布一个或多个命令。该命令可以是写请求或读请求。例如,迁移模块 114 可以向源块设备 108 上的块发布写请求以更新该块。一旦生成了命令,则其被命令驱动器 212 接收到。可以例如使用操作系统环境中的用户空间至内核空间交互来实现命令驱动器 212。

[0055] 命令驱动器 212 接收由迁移模块 114 发布的命令并检查该命令是读请求还是写请求。在该命令是写请求的情况下,命令驱动器 212 使其改向至目的地块设备 110 中的块。在目的地块设备 110 处,执行写命令并更新与源块设备 108 上的块相对应的目的地块设备 110 上的块。在一个实施方式中,命令驱动器 212 基于存储在其他数据 224 中的映射信息来识别用于使写请求改向的块。

[0056] 一旦目的地块设备 110 处的块被更新,则文件系统更新位图 216 中的对应的位也被设定,从而指示更新。在一个实施方式中,由迁移模块 114 来设定文件系统更新位图 216。此外,现在可以在目的地分配位图 222 中将已更新块指示为已分配。在一个实施方式中,迁移模块 114 设置目的地分配位图 222 中的位以在目的地分配位图 222 中指示目的地块设备 110 上的已更新块被分配并具有数据。

[0057] 在一个实施方式中,迁移模块 114 可以将驻留于目的地块设备 110 上的目的地分配位图 222 导入到存储器 204 上并修改目的地分配位图 222。在另一实施方式中,迁移模块 114 可以例如通过在计算设备 112 中提供的(一个或多个)接口 226 来访问并修改目的地块设备 110 上的目的地分配位图 222。

[0058] 返回至由迁移模块 114 生成的请求,情况还可能是命令驱动器 212 可以从迁移模块 114 接收读请求以从源块设备 108 上的(一个或多个)块读取数据。在这种情况下,命令驱动器 212 检查读请求是否被引导到目的地块设备 110 上的已被更新的块。

[0059] 在读请求被引导到在目的地块设备 110 上未被更新的源块设备 108 上的块时,则读请求被引导到源块设备 108 上的块。例如,源块设备 108 上的块可能在目的地块设备 110 上没有被更新,因为该块在源块设备 108 上可能未被粉碎,或者一般地在重新定位中可能未被涉及到。

[0060] 在读请求被引导到在目的地块设备 110 上被更新的块的情况下,命令驱动器 212 将读请求引导到目的地块设备 110 上的对应块。在一个实施方式中,命令驱动器 212 基于文件系统更新位图 216 来识别源块设备 108 处的块在目的地块设备 110 处是否已被更新。在这种情况下,命令驱动器 212 可以基于源块设备 108 和目的地块设备 110 的关联使读请求引导到目的地块设备 110 上的对应块。

[0061] 在一个实施方式中,目的地分配位图 222 还可以指示具有有效数据但未被更新的源块设备 108 上的块。例如,由迁移模块 114 基于源分配位图 220 在目的地分配位图 222 中设置用于源块设备 108 上的此类块的位。此外,目的地分配位图 222 的此类修改可以在目的地块设备 110 上的块已被更新和迁移之后发生。

[0062] 一旦已处理了该命令并在目的地块设备 110 上更新了块,则迁移模块 114 从源块设备 108 拷贝在目的地块设备 110 处未被更新但具有有效数据的那些块。为此,迁移模块 114 确定目的地块设备 110 上的所有此类块的更新状态。例如,迁移模块 114 基于文件系统更新位图 216 来确定块的更新状态。基于块的更新状态,迁移模块 114 将对应的块从源块设备 108 拷贝到目的地块设备 110。例如,迁移模块 114 识别那些块之中的此类块,对于该块而言文件系统更新位图 216 中的对应位被清零,因此指示这些块未被更新。

[0063] 在一个实施方式中,迁移模块 114 基于拷贝位图 218 来拷贝未被更新但具有有效数据的块。可以基于文件系统更新位图 216 和目的地分配位图 222 来计算拷贝位图 218。在一个实施方式中,迁移模块 114 可以通过计算文件系统更新位图 216 和目的地分配位图 222 的 XOR 来计算拷贝位图 218。一旦计算了拷贝位图 218,则迁移模块 114 可以将与在拷贝位图 218 中设置的块相对应的块从源块设备 108 拷贝到目的地块设备 110。

[0064] 可以理解的是虽然文件系统更新位图 216 和拷贝位图 218 被示出为驻留于计算设备 112 上,但两者可以驻留于源设备 102、目的地设备 106、源块设备 108 或目的地块设备 110 的任一上。

[0065] 根据另一实施方式,可以将源块设备 108 上的块分组以形成一个或多个块组。在此类实施方式中,迁移模块 114 可以从源块设备 108 上的块组拷贝块。结合图 3 和图 4 来详细地描述此实施方式。

[0066] 图 3 说明根据本发明的实施方式的示例性文件系统更新位图 216,其包括用于块组的更新位图。在一个实施方式中,文件系统更新位图 216 可以跟踪在目的地块设备 110 处被更新的块和块组。

[0067] 如所指示的,文件系统更新位图 216 包括元素列表 300,其引用一个或多个块组更新位图 302-1、302-2、... 302-N (共同地称为块组更新位图 302)。在一个实施方式中,块组更新位图 302 可以对应于目的地块设备 110 上的块组。元素列表 300 可以包括块组条目 304、更新状态字段 306 以及指针字段 308。块组条目 304 对应于目的地块设备 110 的块组。更新状态字段 306 指示目的地块设备 110 上的块组的更新状态。更新状态字段 306 在被设定时(即 1)指示对应块组中的至少一个块在目的地块设备 110 上已被更新。例如,在块组

更新位图 302-1 中,被设定的位指示目的地块设备 110 上的块组中的对应块已被更新。另一方面,更新状态字段 306 中的清零值(即 0)可以指示对应的块组的块中没有一个被更新。

[0068] 此外,指针字段 308 引用一个或多个块组更新位图 302。如前所述,可以将块组更新位图 302 存储在计算设备 112 的存储器 204 中或源设备 102、目的地设备 106、源块设备 108 或目的地块设备 110 中,并且被元素列表 300 引用。此外,在另一实施方式中,可以将块组更新位图 302 集成在文件系统更新位图 216 中。在一个实施方式中,当迁移模块 114 发布第一写请求以更新块组中的块时,可以为该块组动态地创建块组更新位图 302。

[0069] 已关于块组描述了如图 3 所示的文件系统更新位图 216。在源块设备 108 和目的地块设备 110 具有未被聚集成块组的单独块的情况下,相同的原理将适用。在另一实施方式中,可以用类似于块组位图 302-1 的位图来说明用于文件系统的文件系统更新位图 216,其中,每个设定位将指示目的地块设备 110 上的对应块的已更新状态。

[0070] 图 4 说明根据本发明的一个实施方式的目的地分配位图 222 的示例性结构。可以理解的是可以由类似的结构来说明源分配位图 220。

[0071] 在所述实施方式中,目的地分配位图 222 可以包括一个或多个目的地块组位图 401-1、402-2、... 402-N (共同地称为目的地块组位图 402)。在一个实施方式中,目的地块组位图 402 可以对应于目的地块设备 110 上的块组,并且目的地块组位图 402 中的每个位可以对应于该块组中的块。

[0072] 在目的地块组位图 402 中,设定位指示块组中的对应块具有有效数据且被分配,而清零位指示不具有有效数据的自由块。目的地块组位图 402 中用于块的设定位还可以指示具有有效数据但在目的地块设备 110 处未被更新的源块设备 108 上的对应块。在一个实施方式中,用于在源块设备 108 中的未被更新的块的位例如由迁移模块 114 基于源块设备 108 和目的地块设备 110 的关联在目的地块组位图 402 中被对应地设定。

[0073] 在另一实施方式中,可以不将目的地分配位图 222 分成各个块组。在这种情况下,目的地分配位图 222 可以是表示存储在目的地块设备 110 上的整个文件系统的邻接位图。

[0074] 返回前一情况,其中文件系统的块被布置成块组,例如当迁移模块 114 在目的地块设备 110 上更新对应块组中的块时,修改目的地块组位图 402。在一个实施方式中,可以响应于由迁移模块 114 发布的写请求来实现目的地块设备 110 处的块的更新。例如,利用此类更新,可以使来自源块设备 108 的粉碎块迁移至目的地块设备 110。

[0075] 在块的更新之后,迁移模块 114 识别并拷贝未被更新但具有有效数据的源块设备 108 上的那些块。在一个实施方式中,迁移模块 114 基于拷贝位图 218 来拷贝此类块。

[0076] 迁移模块 114 基于文件系统更新位图 216 来生成拷贝位图 218。在一个实施方式中,迁移模块 114 首先确定块组的更新状态。如果更新状态是无效的,则其指示在该块组内、例如对应于块组条目 304 中的条目 2 的块组内的块中没有一个被更新。在这种情况下,迁移模块 114 拷贝例如由源块组位图 402 所指示的块组内的所有分配块。

[0077] 然而,如果块组的更新状态不是无效的,则其指示块组内的块中的至少一个已被更新。在这种情况下,迁移模块 114 基于块组更新位图 302 和目的地块组位图 402 来产生拷贝位图 218。在这种情况下,迁移模块 114 意图拷贝未被更新但被分配在源块设备 108 上的对应块组中的那些块。在一个实施方式中,迁移模块 114 通过评估块组更新位图 302 中的位和目的地块组位图 402 中的位的 XOR 来生成拷贝位图 218。这样生成的拷贝位图 218

将具有被设定的某些块和被清零的某些块。迁移模块 114 然后将那些块从源块设备 108 拷贝至目的地块设备 110, 对于那些块而言, 对应的位在拷贝位图 218 中被设定。在一个实施方式中, 拷贝位图 218 在文件系统迁移完成时被清零。

[0078] 图 5 和图 6 说明依照本发明的实施例的用于执行文件系统迁移的示例性方法。

[0079] 可以在包含在计算机可读介质上的计算机可执行指令的一般上下文中描述示例性方法。一般地, 计算机可执行指令可以包括例程、程序、对象、组件、数据结构、过程、模块、函数等, 其执行特定功能或实现特定的抽象数据类型。还可以在其中由通过通信网络链接的远程处理设备来执行功能的分布式计算环境中实施所述方法。在分布式计算环境中, 可以将计算机可执行指令定位于本地和远程计算机存储介质(包括存储器存储设备)中。

[0080] 描述所述方法的顺序并不意图被解释为限制, 并且可以按照任何顺序将任何数目的所述方法块组合以实现该方法或替换方法。另外, 在不脱离本文所述的方法、系统以及设备的精神和范围的情况下, 可以从所述方法中删除单独的块。此外, 可以在任何适当的硬件、软件、固件或其组合中实现所述方法。

[0081] 图 5 说明根据本发明的实施方式的处理命令并用于更新文件系统块的示例性方法。在一个实施方式中, 由诸如命令驱动器 212 的命令驱动器来处理该命令。此外, 可以由诸如迁移模块 114 的迁移模块来实现文件系统块的更新。

[0082] 在块 502 处, 使诸如源块设备 108 的源块设备与诸如目的地块设备 110 的目的地块设备相关联。例如, 诸如关联模块 210 的关联模块将源块设备 108 上的块关联并映射到目的地块设备 110 上的对应块。在一个实施方式中, 指示源块设备 108 与目的地块设备 110 之间的映射的信息被存储在诸如其他数据 224 的其他数据中。此外, 一旦实现了源块设备 108 和目的地块设备 110 的映射, 则可以创建更新位图, 诸如文件系统更新位图 216。

[0083] 在块 504 处, 接收用于对块执行读或写操作的一个或多个命令。在一个实施方式中, 迁移模块 114 向源块设备 108 上的块发布命令。该命令可以包括用以向源块设备 108 的块进行读或写的命令。由迁移模块 114 生成的命令被命令驱动器 212 接收。

[0084] 在块 506 处, 确定接收到的命令是否是读请求。例如, 命令驱动器 212 接收该命令并确定从迁移模块 114 接收到的命令是否是读请求。在命令驱动器 212 确定接收到的命令不是读请求的情况下(来自块 506 的“否”路径), 即其为例如意图将块重新定位于源块设备 108 上的写请求, 该命令被引导到目的地块设备 110 上的对应块(块 508)。

[0085] 此外, 可以响应于写请求来更新目的地块设备 108 上的块。在一个实施方式中, 可以通过从源块设备 108 拷贝块并将该块重新定位于目的地块设备 110 中例如以对该块中的数据进行碎片整理来实现该更新。在一个实施方式中, 迁移模块 114 基于源块设备 108 和目的地块设备 110 的关联来识别目的地块设备 110 上的块。在一个实施方式中, 迁移模块 114 生成对应于其中块被更新的块组的块组更新位图, 诸如块组更新位图 302。

[0086] 在命令是读请求的情况下(来自块 506 的“是”路径), 进一步确定(块 510)读请求是否被引导到源块设备 108 上的块, 该块在目的地块设备 110 上已被更新, 例如重新定位在目的地块设备 110 上。在一个实施方式中, 迁移模块 114 基于文件系统更新位图 216 来确定源块设备 108 处的块是否已被更新。

[0087] 如果读请求是针对已被更新的块的(来自块 510 的“是”路径), 则该请求被引导到目的地块设备 110 上的对应块(块 512)。在一个实施方式中, 命令驱动器 212 将读请求引导

到目的地块设备 110。

[0088] 相反,如果在块 510 处确定读请求被引导到在目的地块设备 110 上未被更新的源块设备 108 上的块(来自块 510 的“否”路径),则读请求被引导到源块设备 108 上的块。在一个实施方式中,命令驱动器 212 将该读请求引导到源块设备 108。

[0089] 图 6 说明根据本发明的实施方式的用于执行块的迁移的方法。这些块可以包括例如源块设备 108 上的块,其具有数据且未被更新。在一个实施方式中,可以由迁移模块 114 来执行迁移。

[0090] 在块 602 处,基于文件系统更新位图 216 来确定目的地块设备 110 上的块组的更新状态。在一个实施方式中,迁移模块 114 基于来自文件系统更新位图 216 的更新状态字段(诸如更新状态字段 306)来确定块组的更新状态。如前所述,当块组的块例如响应于由迁移模块 114 发布的写命令在目的地块设备 110 处被更新时,修改文件系统更新位图 216。

[0091] 在另一实施方式中,在块 602 处确定目的地块设备 110 上的块的更新状态。在所述实施方式中,基于文件系统更新位图 216 来确定更新状态。

[0092] 在块 604 处,确定块组的更新状态字段 306 是否被设定。如果在块 604 处用于块组的更新状态字段 306 被设定,即为“1”(来自块 604 的“是”路径),则其指示在目的地块设备 110 上的块组中存在被更新的至少一个块,例如从源块设备 108 被拷贝并在目的地块设备 110 处被进行碎片整理。

[0093] 在另一实施方式中,检查块的更新状态是被设定还是清零。在所述实施方式中,如果状态被设定,即其指示块被更新,则顺次对下一个块再次执行来自块 602 的方法。此外,如果用于块的状态被清零,即其指示块未被更新,则为每个块计算拷贝位图(块 606)。

[0094] 在块 606 处,计算拷贝位图,诸如拷贝位图 218。在一个实施方式中,针对考虑之中的块组生成拷贝位图 218,例如块组拷贝位图。在一个实施方式中,基于文件系统更新位图 216 和目的地分配位图 222 来计算拷贝位图 218。在一个实施方式中,通过确定文件系统更新位图 216 中的块组更新位图 302 和目的地分配位图 222 中的目的地块组位图 402 的 XOR 来生成拷贝位图 218。拷贝位图 218 可以指示将从源块设备 108 拷贝的块。在一个实施方式中,块组拷贝位图中的设定位说明在目的地块设备 110 处未被更新但具有数据的源块设备 108 中的那些块。

[0095] 在另一实施方式中,针对文件系统更新位图 216 中的其更新状态被清零的块单独地生成拷贝位图 218。在所述实施方式中,基于文件系统更新位图 216 和目的地分配位图 222 来计算拷贝位图 218。例如,在这种情况下,通过计算与文件系统更新位图 216 中的其更新状态被清零的那些块相对应的目的地分配位图 222 中的位和文件系统更新位图 216 中的位的 XOR 来生成拷贝位图 218。

[0096] 在块 608 处,基于拷贝位图 218 将来自源块设备 108 的块拷贝到目的地块设备 110。在一个实施方式中,将与拷贝位图 218 中的设定位相对应的块从源块设备 108 拷贝到目的地块设备 110。

[0097] 返回至判定块 604,如果更新状态字段 306 未被设定,即文件系统更新位图 216 中的更新状态字段 306 是“0”(来自块 604 的“否”路径),则具有数据的来自源块设备 108 的块组的块在块 610 处被拷贝到目的地块设备 110 中的对应块。在一个实施方式中,块的拷贝基于源分配位图,诸如源分配位图 220。例如,基于源分配位图 220 中的诸如源块组位图

402 的源块组位图来从源块设备 108 拷贝块。在一个实施方式中,源分配位图 220 指示在源块设备 108 上具有数据的那些块。例如,源分配位图 220 中的设定位可以指示在源块设备 108 上具有数据的块,并且在块 610 处从源块设备 108 拷贝此类块。

[0098] 本文所述的方法和设备在动态重新定位目的地块设备上的文件系统块的情况下执行文件系统迁移而不影响源块设备上的文件系统。根据一方面,目的地块设备上的文件系统块的重新定位可以缩小文件系统的大小。利用此类文件系统迁移,可以在没有任何数据损失的情况下使文件系统迁移至小于源块设备的目的地块设备。此外,可以采用所述方法和设备来迁移碎片化的文件系统以在目的地块设备处实现有效的存储合并、高效的盘空间利用 / 优化。

[0099] 结论

虽然用结构特征和 / 或方法所特有的语言描述了用于执行文件系统迁移的实施方式,但应理解的是本发明(以及未附的权利要求)不一定局限于所述的特定特征或方法。相反,用于执行文件系统迁移的特定特征和方法是作为本发明的示例性实施方式而公开的。

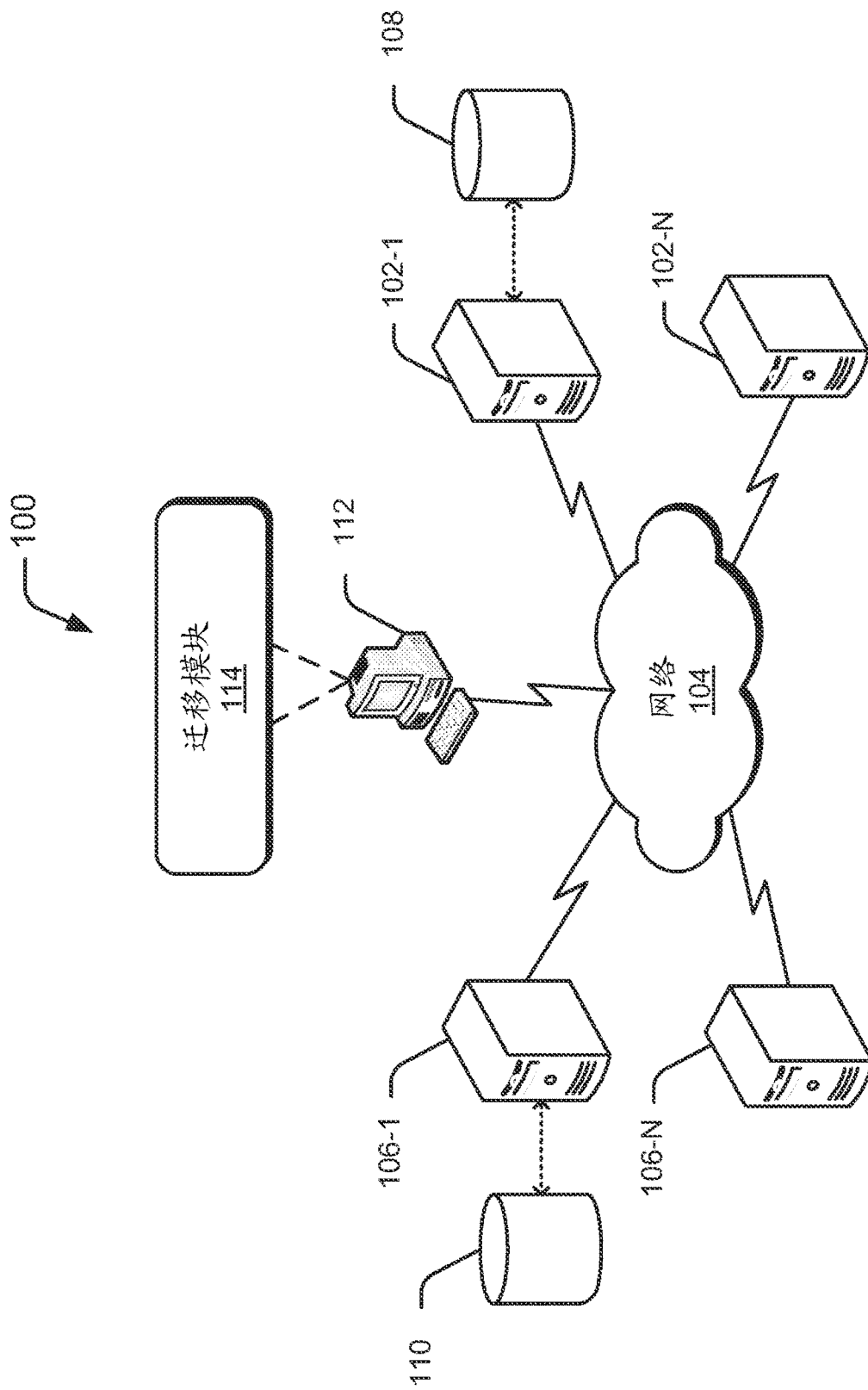


图 1



图 2

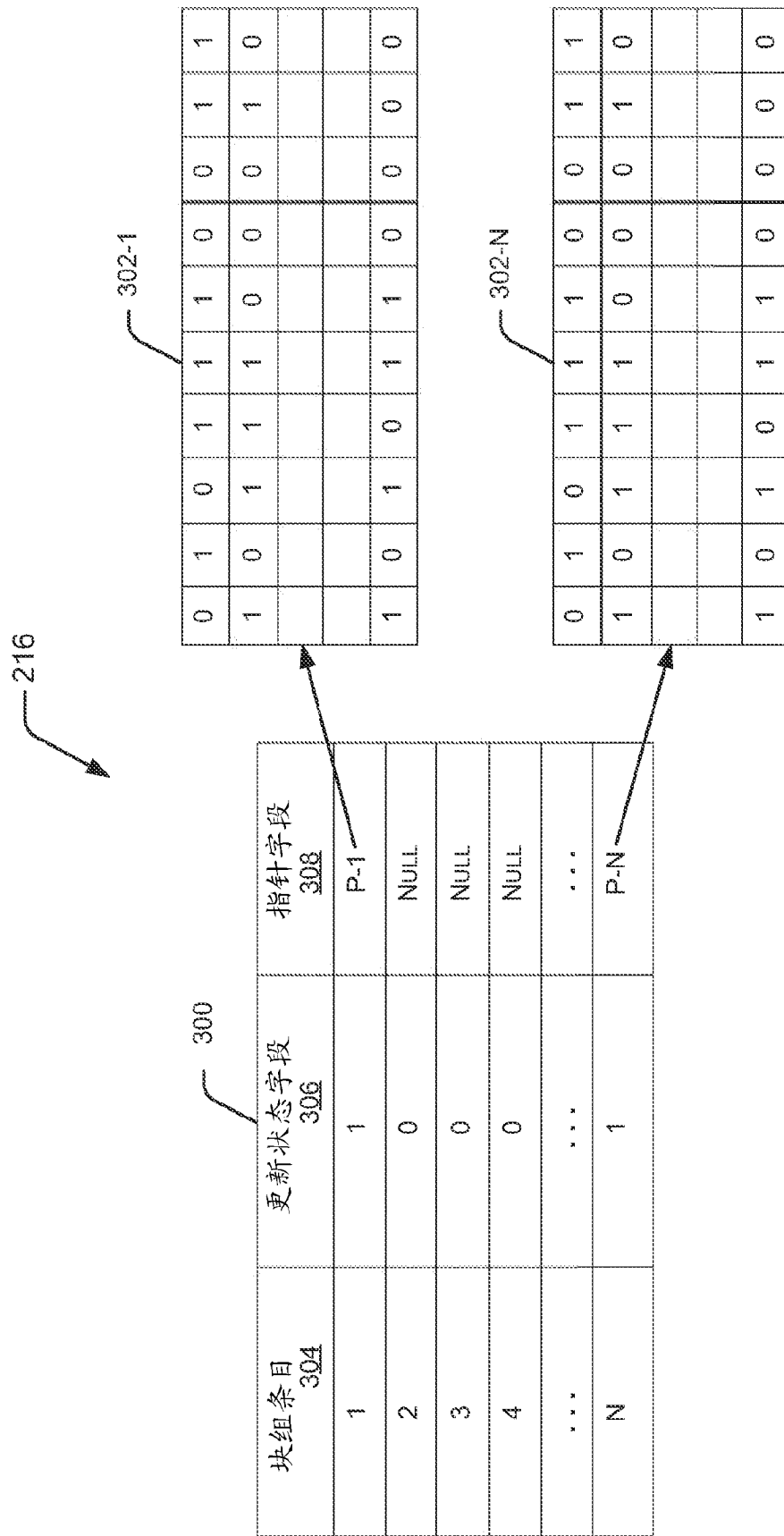


图 3

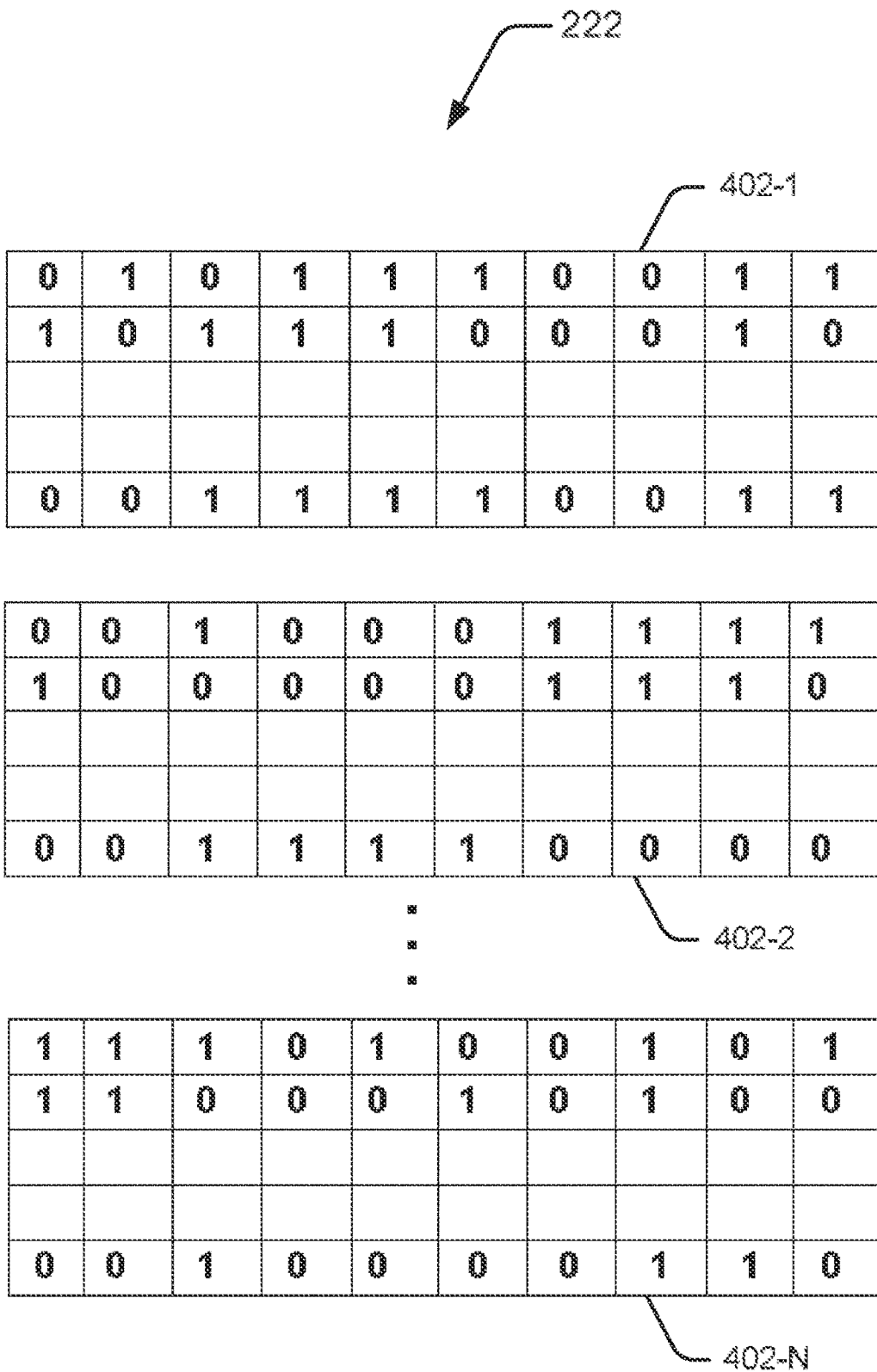


图 4

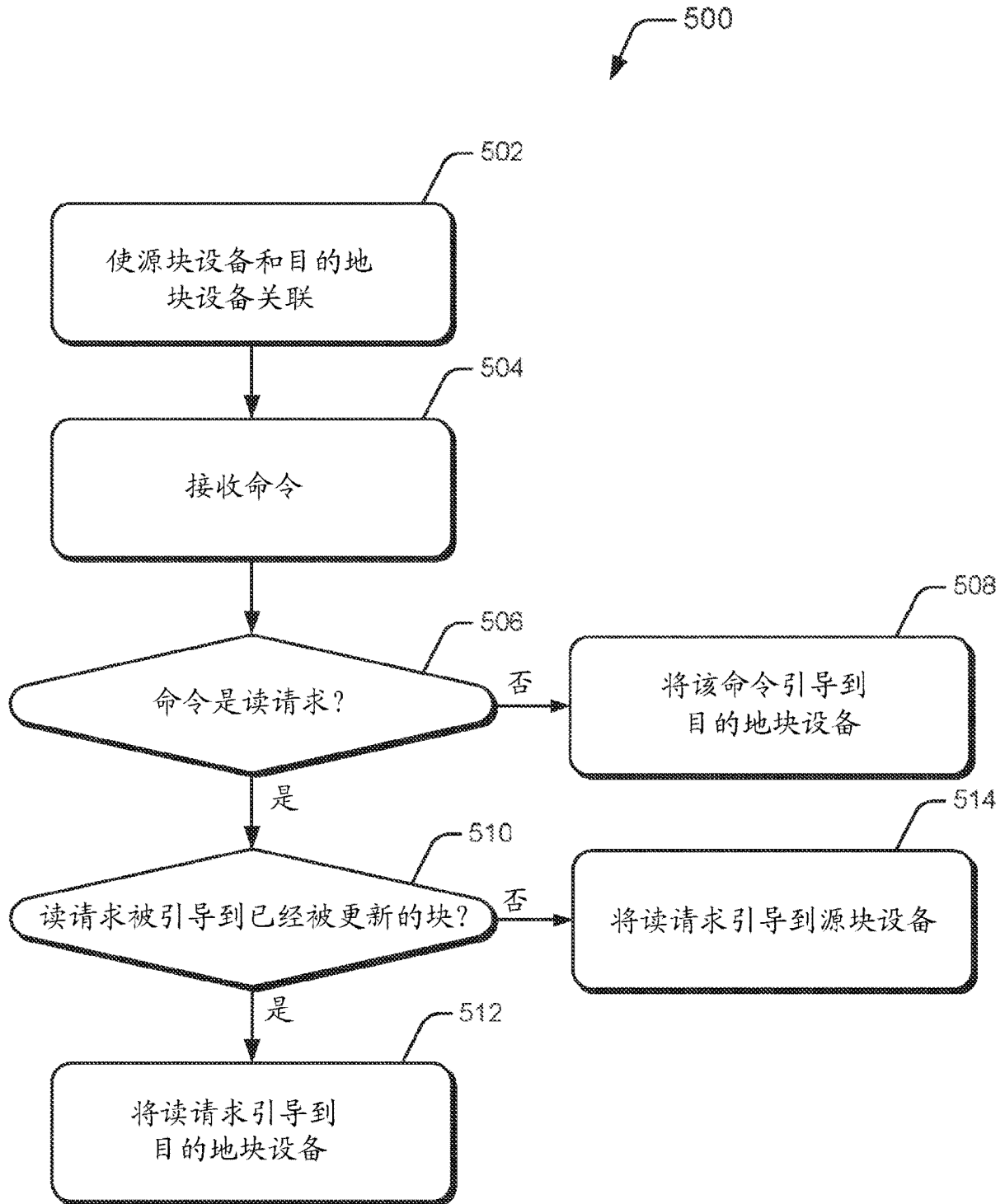


图 5

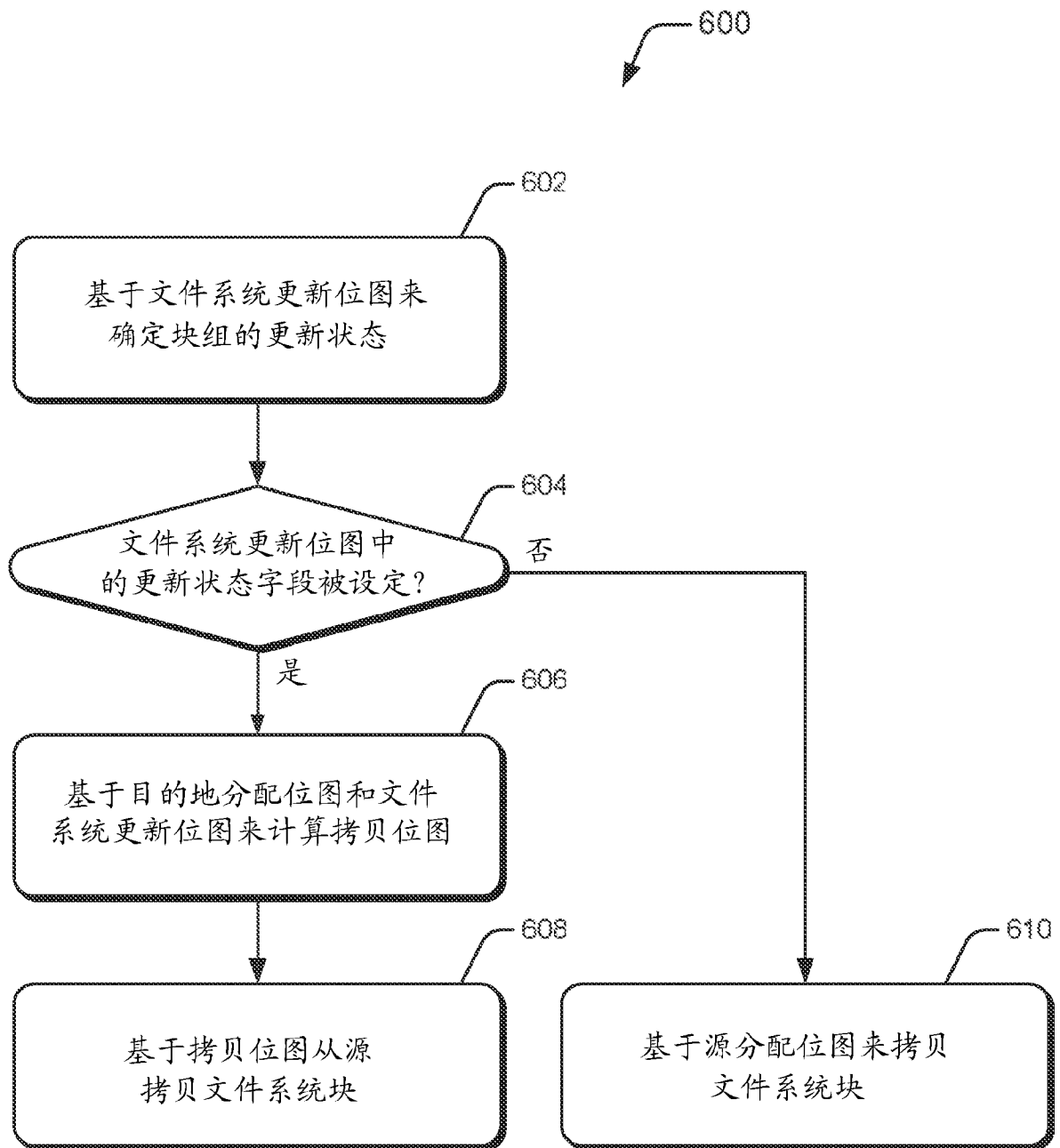


图 6