



(12)发明专利

(10)授权公告号 CN 104272244 B

(45)授权公告日 2017.09.29

(21)申请号 201380023646.5

(22)申请日 2013.03.11

(65)同一申请的已公布的文献号
申请公布号 CN 104272244 A

(43)申请公布日 2015.01.07

(30)优先权数据
13/436,554 2012.03.30 US

(85)PCT国际申请进入国家阶段日
2014.11.05

(86)PCT国际申请的申请数据
PCT/US2013/030279 2013.03.11

(87)PCT国际申请的公布数据
W02013/148143 EN 2013.10.03

(73)专利权人 NETAPP股份有限公司

地址 美国加利福尼亚

(72)发明人 V·K·达盖 C·约翰斯顿 郑玲

(74)专利代理机构 北京市金杜律师事务所
11256

代理人 鄢迅

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

(56)对比文件
US 8095756 B1,2012.01.10,
CN 101916171 A,2010.12.15,
US 2011225214 A1,2011.09.15,

审查员 李中兴

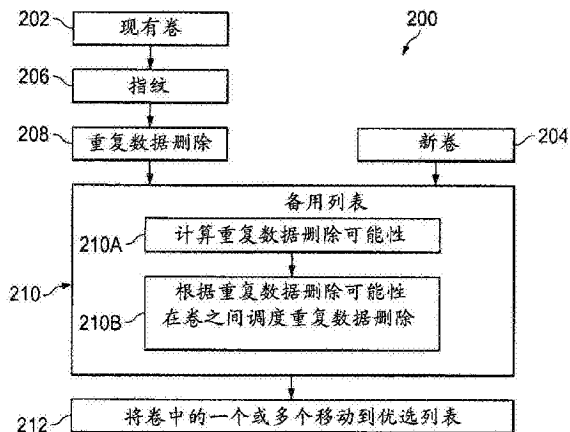
权利要求书3页 说明书10页 附图3页

(54)发明名称

用于对处理进行调度以实现空间节省的系统、方法

(57)摘要

在具有存储到存储硬件的多个卷的系统中执行的方法,该方法包括:对于卷中的每一个,随着时间而迭代地产生各空间节省可能性;和通过对于空间节省可能性分析卷中的每一个并且至少部分基于空间节省可能性指派资源的优先级来在多个卷之间调度空间节省操作。



1. 一种在具有存储到存储硬件的多个存储对象的系统中执行的方法,所述方法包括:
将所述多个存储对象指派到备用列表;
对于所述备用列表中的所述多个存储对象中的每一个存储对象,产生相应空间节省可能性;
基于与所述多个存储对象的第一子集对应的所产生的空间节省可能性,将所述多个存储对象的所述第一子集从所述备用列表重新指派到优选列表;
基于与所述多个存储对象的第二子集对应的所产生的空间节省可能性,将所述多个存储对象的所述第二子集从所述备用列表重新指派到第三列表;以及
通过分析所述多个存储对象中的每一个存储对象并且至少部分基于所述多个存储对象的所产生的空间节省可能性指派资源的优先级来在所述多个存储对象之间调度空间节省操作,其中所述调度对这样的空间节省操作进行调度:所述空间节省操作在对被指派到所述备用列表的任何存储对象执行之前要对被指派到所述优选列表的存储对象执行,并且不调度要对被指派到所述第三列表的所述存储对象执行的所述空间节省操作。
2. 根据权利要求1所述的方法,其中,产生相应空间节省可能性包括:
对于所述多个存储对象中的每一个存储对象:
识别自先前的空间节省操作以来接收的数据;
对所识别的数据执行另一空间节省操作,并且对于所识别的数据计算存储空间中的当前百分比减小;以及
执行所识别的数据的所述当前百分比减小与对于先前添加的数据的累积百分比减小的加权平均,其中,所述加权平均产生所述相应空间节省可能性的所述百分比减小。
3. 根据权利要求1所述的方法,其中,对于给定存储对象的相应空间节省可能性基于对于所述给定存储对象的先前的空间节省结果。
4. 根据权利要求1所述的方法,其中,调度包括:
识别所述存储对象中的两个或更多个存储对象具有类似的空间节省可能性;以及
响应于所述识别,基于相对改变日志大小来在所述存储对象中的所述两个或更多个存储对象之间调度空间节省操作。
5. 根据权利要求1所述的方法,其中,将所述多个存储对象的所述第一子集重新指派到所述优选列表至少部分基于所述第一子集中的所述存储对象是否具有跟上数据改变速率的空间节省操作。
6. 根据权利要求1所述的方法,进一步包括:
响应于系统资源的变化或者与所述存储对象对应的所述空间节省可能性的变化,将所述多个存储对象的所述第一子集中的存储对象从所述优选列表重新指派到所述备用列表。
7. 根据权利要求1所述的方法,其中,所述优选列表包括具有指示比在所述备用列表上的每一个存储对象更大的空间节省的相应空间节省可能性的所述存储对象中的存储对象。
8. 根据权利要求7所述的方法,其中,将所述多个存储对象的所述第二子集指派到所述第三列表至少部分基于所述第二子集中的所述存储对象是否具有指示比在所述备用列表上的每一个存储对象更低的空间节省的相应空间节省可能性。
9. 根据权利要求1所述的方法,所述方法由网络存储系统中的存储服务器执行。
10. 一种用于对于多个存储对象管理空间节省的装置,所述装置包括:

用于对所述多个存储对象随着时间而执行多个空间节省操作的部件；

用于计算对于所述多个存储对象中的每一个存储对象的相应空间节省值以使得对于给定存储对象的每个随后的空间节省操作导致对于所述给定存储对象的所述空间节省值的迭代重新计算的部件；以及

用于通过以下各项来将处理资源分配给空间节省请求的部件：

将所述多个存储对象指派到备用列表；

基于与所述多个存储对象的第一子集对应的所计算的空间节省值，将所述第一子集从所述备用列表重新指派到优选列表；

基于与所述多个存储对象的第二子集对应的所计算的空间节省值，将所述第二子集从所述备用列表重新指派到第三列表；以及

通过对所述多个空间节省操作中的这样的空间节省操作进行调度：所述空间节省操作在对被指派到所述备用列表的任何存储对象执行之前要对被指派到所述优选列表的存储对象执行，并且不调度要对被指派到所述第三列表的所述存储对象执行的所述空间节省操作，来在所述多个存储对象之间调度空间节省请求。

11. 根据权利要求10所述的装置，其中，将所述多个存储对象的所述第一子集重新指派到所述优选列表至少部分基于所述第一子集中的卷是否具有跟上数据改变速率的空间节省操作。

12. 根据权利要求10所述的装置，进一步包括：

用于响应于系统资源的变化或者与所述存储对象对应的所述空间节省可能性的变化，将所述多个存储对象的所述第一子集中的存储对象从所述优选列表重新指派到所述备用列表的部件。

13. 根据权利要求10所述的装置，其中，所述优选列表包括具有指示比所述备用列表上的每一个存储对象更大的空间节省的相应空间节省可能性的所述存储对象中的存储对象。

14. 一种具有存储到存储硬件的多个存储对象的系统，包括：

存储子系统，所述存储子系统包括非易失性存储介质，所述非易失性存储介质被虚拟地划分成多个存储对象；以及

计算机，所述计算机与所述存储子系统通信，所述存储子系统管理对于所述多个存储对象内的空间节省的处理资源的分配，所述计算机被配置为执行以下动作：

通过多个空间节省操作迭代地计算对于所述多个存储对象中的每一个存储对象的相应空间节省可能性；以及

通过以下各项对于空间节省在所述多个存储对象之间进行选择：

将所述多个存储对象指派到备用列表；

基于与所述多个存储对象的第一子集对应的所计算的空间节省可能性，将所述第一子集从所述备用列表重新指派到优选列表；

基于与所述多个存储对象的第二子集对应的所计算的空间节省可能性，将所述第二子集从所述备用列表重新指派到第三列表；以及

通过对这样的空间节省操作进行调度：所述空间节省操作在对被指派到所述备用列表的任何存储对象执行之前要对被指派到所述优选列表的存储对象执行，并且不调度要对被指派到所述第三列表的所述存储对象执行的所述空间节省操作，来在所述多个存储对象之

间调度所述空间节省操作。

15. 根据权利要求14所述的系统,其中,所述计算机包括通过存储网络与所述存储子系统进行通信的存储服务器。

16. 根据权利要求14所述的系统,其中,所述第一子集具有相对于所述第二子集的相应空间节省可能性更高的相应空间节省可能性。

17. 根据权利要求14所述的系统,其中,所述相应空间节省可能性通过执行包括先前的空间节省操作的结果的加权平均来计算。

18. 根据权利要求14所述的系统,其中,所述相应空间节省可能性包括重复数据删除可能性和压缩可能性中的至少一个。

用于对处理进行调度以实现空间节省的系统、方法

技术领域

[0001] 本说明书一般涉及网络存储系统中的空间节省操作(例如,压缩和重复数据删除),更具体地,涉及用于在重复数据删除处理中在卷之间分配空间节省资源的技术。

背景技术

[0002] 许多企业包括连接到客户端计算系统的网络存储系统(诸如网络附属存储(NAS)和存储区域网络(SAN)),由此客户端可以访问由存储系统管理的数据。从用户的(例如,客户端的)观点来讲,网络存储系统可以包括一个或多个存储对象(存储卷),这些存储对象通常被称为逻辑卷或虚拟卷。这样的网络存储系统可以存储非常大量的重复数据,因此,在一些情况下执行重复数据删除以便更高效地使用可用的存储空间可能是希望的。在数据可以在网络存储系统中进行重复数据删除的程度上,重复数据的移除在一些情况下可以提供大量的存储空间节省,由此可能省钱。

[0003] 用于网络存储的一些常规技术实现与各虚拟卷对应的文件系统,该文件系统提供在虚拟卷内逻辑组织的较低级存储容器(例如,文件)的分层组织,并且利用指针来指向底层数据,其中,底层数据以数据块布置。给定文件可以指向多个块,并且一个块可以与多个文件相关联。此外,给定文件可以包括在另一个文件中重复的数据。例如,存储卷可以包括多个电子邮件收件箱,每个收件箱包括特定的电子邮件附件。在大多数情况下,存储电子邮件附件的多个副本将是不希望的,因为这样做将浪费存储资源。一些常规的重复数据删除操作通过仅留下数据的单个副本并且利用对于该单个副本的指针取代重复副本来避免保存数据片的多个副本。因此,多个文件与同一数据相关联,但是避免了该数据的重复副本。

[0004] 重复数据删除操作可能使用大量的处理资源。在一个示例中,常规的重复数据删除处理在尚未进行重复数据删除的卷上开始。常规的重复数据删除处理包括:从存储器(通常为硬盘)读取数据块,创建用于数据块中的每一个的指纹(例如,指纹可以是指示块中的数据的小数据片),以及比较指纹以确定哪些块可能为重复。重复数据然后如上所述用指针取代。然而,通常,该处理可能使用大量的处理能力,从用户的角度来讲,这可以将其自身表现为降低的性能。此外,从盘读取大量的数据块可能会花费相对长的时间。

[0005] 另一种常规的重复数据删除处理在数据块被保存或者从一个卷传送到另一个卷(例如,在备份操作中)时产生该数据块的指纹。与以上示例对比,这应消除在单个操作中读取整个卷以对该整个卷取指纹。然而,仅仅将指纹彼此进行比较就可能使用大量的处理资源。

[0006] 简言之,从用户的观点来讲,由于被分配到重复数据删除处理而没有用于对于用户更可见的并发存储和检索操作的处理资源,重复数据删除处理可能导致感觉到性能的不足。假定网络存储系统具有有限量的处理资源以专用于它执行的各种操作,则高效地执行重复数据删除以便从最小量的处理资源得到大量的重复数据删除将是希望的。

附图说明

[0007] 当与附图一起阅读时,从以下详细描述最好地理解本公开。

[0008] 图1是实现可以实现各种实施例的存储操作系统的示例网络存储系统的图示。

[0009] 图2是根据一个实施例改动的在众多卷之间调度重复数据删除操作的示例方法的图示。

[0010] 图3是根据一个实施例改动的可与图2中所示的方法一起操作的用于在众多卷之间分配用于重复数据删除的资源的方法的图示。

[0011] 图4是根据实施例改动的可与图2中所示的方法一起操作的在卷 之间分配重复数据删除资源的示例方法的图示。

发明内容

[0012] 各种实施例包括提供分配重复数据删除资源的高效方法的系统、方法和计算机程序产品。在一个示例中,对于多个存储对象执行多个重复数据删除操作,以使得存储对象中的每一个的传入数据不时地被重复数据删除为现有数据。对于存储对象中的每一个,基于重复数据删除操作的结果来计算重复数据删除值。重复数据删除值提供对于存储对象中的每一个随着时间而实现的重复数据删除量的指示。在这个示例中,假定对于存储对象的过去结果是期望的未来性能的指示,因此,重复数据删除值是重复数据删除节省可能性(potential)的指示。在随后的重复数据删除操作中,具有较高的重复数据删除可能性的存储对象被选择以在具有较低的可能性的其它存储对象之前进行重复数据删除。随着时间而显示出较高的重复数据删除可能性的存储对象可能最终被添加到优选列表。以这种方式,对于空间节省具有较高的可能性的存储对象被给予重复数据删除资源的优先级。

[0013] 本公开的更广泛的形式的一种涉及在具有存储到存储硬件的多个存储对象的系统中执行的方法,该方法包括:对于存储对象中的每一个,随着时间而迭代地产生各空间节省可能性;和通过对于空间节省可能性分析存储对象中的每一个并且至少部分基于空间节省可能性指派资源的优先级来在多个存储对象之间调度空间节省操作。

[0014] 本公开的更广泛的形式的一种涉及具有计算机可读介质的计算机程序产品,该计算机可读介质有形地记录用于对于多个存储对象管理空间节省的计算机程序逻辑,该计算机程序产品包括:对多个存储对象随着时间而执行多个空间节省操作的代码;计算存储对象中的每一个的各空间节省值以使得对于给定存储对象的每个随后的空间节省操作导致该给定存储对象的空间节省值的迭代重新计算的代码;以及通过对于具有较高的各空间节省值的存储对象的子集优先考虑处理资源来将该处理资源分配给空间节省请求的代码。

[0015] 本公开的更广泛的形式的一种涉及包括存储子系统和计算机的系统,该存储子系统包括非易失性存储介质,该非易失性存储介质被虚拟地划分成多个存储对象;该计算机与存储子系统通信,管理多个存储对象内的空间节省的处理资源的分配,该计算机被配置为执行以下动作:关于多个空间节省操作迭代地计算存储对象中的每一个的各空间节省可能性;和对于空间节省在存储对象之间进行选择,以使得具有比第二存储对象的第二空间节省可能性高的第一空间节省可能性的第一存储对象在至少部分基于第一空间节省可能性与第二空间节省可能性之间的差异来分配用于空间节省的处理资源时接收到优先。

具体实施方式

[0016] 以下公开提供用于实现本发明的不同特征的许多不同的实施例或示例。以下描述组件和布置的具体示例以简化本公开。当然,这些仅仅是示例,而非意图限制。另外,本公开可能在各个示例中重复引用数字和/或字母。该重复是出于简化和清晰的目的,并非本身决定所讨论的各种实施例和/或配置之间的关系。

[0017] 理解的是,各种实施例可以在任意类型的网络环境中实现。例如,一些实施例可以在连接多个客户端和多个服务器的TCP/IP网络中实现。其它实施例可以在更专用的环境中实现,诸如在网络附属存储(NAS)、存储区域网络(SAN)或任意其它网络存储配置中实现。各种实施例不限于任意特定的网络架构或目的,并且可以进行改动以供在范围广泛的一系列环境中使用。

[0018] 本文中公开的各种实施例提供根据给定卷的所计算的可能性来分配重复数据删除资源的技术。具有较高的重复数据删除可能性的卷因此更有可能进行重复数据删除,而具有较低的重复数据删除可能性的卷不太可能进行重复数据删除。

[0019] 在一些系统中,用于重复数据删除的总工作负荷可能大于系统中可以分配给重复数据删除的处理资源的量。例如,重复数据删除在一些应用中可能是处理密集的,并且系统可能仅运行重复数据删除操作,只要它们不使存储和检索功能降低超出客户的性能期望。因此,由于系统的有限的处理资源,一些卷可能放弃重复数据删除。

[0020] 本文中描述的示例实施例通过在最有可能看见存储空间节省的益处的卷上使用处理资源来提供对于有限的这些处理资源的问题的解决方案。在一个示例中,各个卷在第一列表中被分组,在这个示例中第一列表也被称为备用列表。随着时间继续,可以对卷进行重复数据删除,并且对于每个卷计算重复数据删除值以提供该卷的重复数据删除可能性的指示。在一些情况下,使用通过重复数据删除而对于给定卷节省的空间量来计算重复数据删除值。对于卷中的每一个可以不时地对传入数据执行重复数据删除,并且迭代地计算重复数据删除值,使得后续的重复数据删除操作根据通过该重复数据删除操作实现的空间节省来影响卷的重复数据删除值。从一个角度来看,重复数据删除值是后顾性的,因为它是使用在过去实现的重复数据删除结果(例如,空间节省)来计算的。重复数据删除值也是前瞻性的,因为在这个示例中假定过去的重复数据删除结果是未来的节省的指示。

[0021] 随着时间的过去,卷将具有与每一个卷相关联的不同的重复数据删除值。如以上所提到的,运行重复数据删除应用的设备(例如,网络存储服务器)将更经常地选择具有较高的节省可能性的卷以进行重复数据删除。一致地显示出较高的节省可能性的一些卷最终可以被移动到优选列表,在优选列表中,这些卷在对备用列表中的任一卷进行重复数据删除之前进行重复数据删除。因此,这个示例实施例通过选择卷中展现出较高的空间节省可能性的一些卷来填充优选列表。在一些情况下,可以将显示出最小节省可能性的卷移动到第三列表,其将这些卷从重复数据删除的考虑中移除。

[0022] 类似的处理可以用于其它空间节省技术,诸如压缩等。例如,一些实施例可以随着时间而迭代地计算压缩的空间节省值,并然后调度压缩处理资源,以使得具有较高的空间节省可能性的卷在其它卷之前被压缩。

[0023] 图1是实现可以实现各种实施例的存储操作系统(未示出)的示例网络存储系统100的图示。存储服务器102通过网络103耦合到一组客户端101。网络103可以包括例如局域网(LAN)、广域网(WAN)、互联网、光纤通道架构或这样的互连的任意组合。服务器102、主

存储子系统104和次存储子系统106可以连接到共同的通信和数据传送基础设施,诸如光纤通道,并且可以共同包括SAN或NAS。

[0024] 客户端101中的每一个可以包括例如个人计算机(PC)、传统的服务器计算机、工作站、手持计算/通信设备或平板等。图1示出了三个客户端101a-c,但是实施例的范围可以包括任意适当数量的客户端。

[0025] 在一些实施例中,客户端101中的一个或多个可以用作管理站。这样的客户端可以包括管理应用软件,该管理应用软件被网络管理员使用以配置存储服务器102、对主存储子系统104中的储存器进行配给、以及执行与存储网络相关的其它管理功能,诸如调度备份、设置用户访问权限等。

[0026] 存储服务器102管理主存储子系统104中的数据存储。存储服务器102对来自客户端101的读和写请求进行处理,其中,该请求针对存储在或者要被存储在主存储子系统104中的数据。主存储子系统104不限于任意特定的存储技术,并且可以使用现在已知或者以后开发的任意存储技术。例如,主存储子系统104具有若干个非易失性大容量存储设备(未示出),该非易失性大容量存储设备可以包括常规的磁盘或光盘或磁带驱动器、非易失性固态存储器(诸如闪存)或其任意组合。在一个特定示例中,主存储子系统104可以包括一个或多个独立磁盘冗余阵列(RAID)。在这个示例中,主存储子系统104包括这里由主卷105表示的多个虚拟卷。

[0027] 类似地,存储服务器102对次存储子系统106处的备份操作进行处理。次存储子系统106不限于任意特定的存储技术,而是可以使用常规的磁盘或光盘或磁带驱动器、非易失性固态存储器(诸如闪存)、一个或多个RAID、或其任意组合。次卷107可以包括多个虚拟卷。

[0028] 这个示例中的存储服务器102对卷105和/或对卷107提供重复数据删除操作。具体地,存储服务器102包括空间节省管理器应用110,该空间节省管理器应用根据图2-4的方法执行指令以对数据进行重复数据删除和/或压缩。

[0029] 存储服务器102可以允许根据任意适当的协议或存储环境配置的数据访问。在一个示例中,存储服务器102如在NAS环境中常规执行的那样对客户端101提供文件级数据访问服务。在另一个示例中,存储服务器102如在SAN环境中常规执行的那样提供块级数据访问服务。在又一个示例中,存储服务器102对客户端101提供文件级数据访问服务和块级数据访问服务两者。

[0030] 在一些示例中,存储服务器102具有分布式架构。例如,一些实施例中的存储服务器102可以被设计为物理上分离的网络模块(例如,“N-blade”)和数据模块(例如,“D-blade”),该网络模块和数据模块彼此通过物理互连进行通信。为促成分布式架构,空间节省管理器110的一个或多个功能可以由次存储子系统106处和/或主存储子系统104处的硬件和/或软件实现。存储操作系统在服务器102上运行,并且提供空间节省管理器110,并且在各种实施例中,存储子系统104和106可以包括与存储操作系统兼容的专用硬件和/或软件以实现空间节省。

[0031] 系统100仅作为示例示出。其它类型的硬件和软件配置可以根据本文中描述的特征进行改动以供使用。例如,各种实施例可以在具有多个卷的任意系统(不管是否基于网络)中实现。

[0032] 图2是根据一个实施例改动的示例性方法200的图示。方法200可以由图1的空间节

省管理器110执行以对众多的存储卷进行重复数据删除和/或压缩。提供图2以阐明空间节省管理器可以如何从被指派要控制的(现有的或新的)卷的初始点开始进行空间节省处理。图2不排除其它卷已经进行了重复数据删除并且具有所计算的重复数据删除值的可能性。

[0033] 图2的示例集中于重复数据删除作为空间节省技术,并且理解的是,实施例的范围包括任意适当的空间节省技术。例如,一些实施例可以使用压缩,并且以下描述的原理如应用于重复数据删除那样也应用于压缩。

[0034] 在动作202中,空间节省管理器(在这个示例中,重复数据删除管理器)对现有卷开始处理。例如,现有卷可以是已经具有多于微不足道的量的数据存储到其的卷。

[0035] 在动作204中,重复数据删除管理器对新卷开始处理。例如,新卷可以是尚未具有很少的数据存储到其或者没有数据存储到其的卷。对现有卷开始处理(动作202)与对新卷开始处理(动作204)相比之间的不同在于,现有卷已经具有要进行取指纹和重复数据删除的数据,而新卷则可以跳过取指纹和重复数据删除,并直接进入改变日志记录(以下对其进行进一步描述)。

[0036] 在动作206中,重复数据删除管理器对于现有卷中的数据产生指纹。如以上所提到的,一些实施例包括将数据以被称为数据块的固定大小的片保存在卷上。产生指纹的一个示例包括使用诸如消息摘要算法5(MD5)的算法对块进行哈希以产生比特大小比原始数据块小得多、但是仍然提供数据块的内容的指示的数。MD5仅是示例性的,因为各种实施例可以使用任意适当的取指纹算法。

[0037] 例如,一些取指纹算法可以对于比块的大小小或者比块的大小大的数据单元产生指纹。实际上,一些指纹算法具有可变长度的能力,以便可适应不同于数据块边界的数据片。而且,一些实施例可以利用强的指纹,使得不同数据块具有相同指纹的可能非常低。在权衡需要较少的处理以产生较弱的指纹的情况下,其它实施例可以利用较弱的指纹,接受两个不同数据块具有相同指纹的稍微大的风险。实施例的范围不限于产生指纹的任意特定处理。

[0038] 此外,动作206可以包括:从作为存储卷的基础的物理存储介质读取数据,并然后在数据被读取之后从该数据产生指纹。如以上所提到的,在一些系统中,从物理存储介质读取数据可以是处理密集的。动作206还可以包括将指纹保存到指纹数据库。

[0039] 在动作208中,重复数据删除管理器使用指纹数据库中的指纹对现有卷执行重复数据删除。指纹对于底层数据被用作代理,其彼此进行比较以识别重复的指纹。为了识别重复的指纹,重复数据删除管理器可以按照升序或降序的指纹对指纹数据库进行排序。例如,每个指纹可以包括字母数字哈希值(或布尔哈希值或任意其它适当的哈希值),并且重复数据删除引擎可以按照哈希值进行排序。重复由相同的哈希指示,该相同的哈希在排序之后是相邻的。如果找到重复的指纹,则重复数据删除管理器从卷中检索实际的数据块,并且执行该数据块的逐字节比较。如果逐字节比较验证该数据块为重复,则重复数据删除引擎利用对于单个保留副本的指针取代重复数据块中的每一个,除了一个之外。在一些情况下,指纹可以足够强以使得可以省略可疑的重复块的逐字节比较,使得可以在没有进一步验证的情况下对通过它们的指纹指示为重复的块进行重复数据删除。

[0040] 动作208还可以包括对于现有卷计算重复数据删除值。例如,如果重复数据删除处理从1太字节(TB)的卷消除了200千兆字节(GB),则对于第一个TB,节省可以为20%。

[0041] 在动作210中,重复数据删除引擎将现有卷指派给备用列表,并且还将新卷指派给备用列表。如以上所提到的,方法200不排除已经在处理中的另外的卷,并且其它卷可以另外被指派给备用列表。并且,其它卷可能已经被指派给优选列表(以下对其进一步解释)。

[0042] 重复数据删除管理器对备用列表上的卷执行改变日志记录。在这个示例中,改变日志记录包括当数据被写入到特定卷时创建存储到该卷的数据的指纹。改变日志表示自最近的重复数据删除处理以来对于卷的数据改变。卷的改变日志随着自最后的重复数据删除处理以来数据被写入到该卷而增长。在下一次重复数据删除处理期间,重复数据删除管理器将改变日志中的该指纹与改变日志中的其它指纹进行比较以确定重复数据是否已被添加。

[0043] 重复数据删除管理器还可以将改变日志与指纹数据库进行比较。如以上所解释的,当找到重复的指纹时,在仅留下单个保留副本并用对于单个保留块的引用(例如,指针)取代块的其余部分之前,可以将或者可以不将底层数据块验证为重复。重复的指纹被删除,并且改变日志的剩余指纹被添加到指纹数据库,所以,指纹数据库随着每次重复数据删除操作而增长。类似地,改变日志在重复数据删除操作之间增长,但是在重复数据删除操作完成时返回到零或接近零。

[0044] 动作210a包括对于备用列表上的卷中的每一个计算重复数据删除可能性。通过给定卷的每次重复数据删除操作,该卷的重复数据删除值被更新以说明通过最近的重复数据删除操作实现的空间节省。在许多情况下,重复数据删除值的这个迭代计算包括对于旧的空间节省和新的空间节省使用加权平均。在以上的一个示例中,对于第一个TB的20%的重复数据删除值被计算。继续这个示例,如果随后的重复数据删除操作对于100GB的数据导致10%的空间节省,则与对10%节省进行加权相比,在计算中,该计算可以对20%节省进行十倍的加权(因为1TB为100GB的十倍)。对于1.1TB,所得的重复数据删除值为19.1%。对于每次重复数据删除操作,对于每个卷迭代地计算重复数据删除值。

[0045] 在一个示例中,重复数据删除管理器在它更新节省可能性时记录节省可能性应用的用户数据量。在随后的操作之后,重复数据删除管理器利用该量对旧的节省可能性进行加权,并且与通过在操作中处理的数据量进行加权的新的节省可能性相加。等式(1)提供了迭代地计算节省可能性的一种示例技术。

[0046] Eq. (1) 新SP = (旧SP*旧卷大小) + (操作SP*操作处理的数据) / (旧卷大小+操作处理的数据),其中,SP是节省可能性,新卷大小=卷中的用户数据量。空间节省管理器可以使用任意适当的数据结构来存储这样的信息。

[0047] 卷中的每一个在它处于操作中时连续地存储数据,从而相应地使其改变日志增加。在一个示例中,当改变日志达到预定容量阈值(例如,为20%整)时,重复数据删除管理器将重复数据删除请求输入到其对于该特定卷的重复数据删除队列中。如以下在动作210b处所解释的,重复数据删除队列中的请求不一定以它们被输入的次序完成。相反,以上对于备用日志中的卷描述的重复数据删除操作根据给定卷的重复数据删除可能性来分配。

[0048] 在动作210b处,重复数据删除管理器根据重复数据删除可能性来在备用列表卷之间调度重复数据删除。例如,在对备用列表卷进行重复数据删除之前,重复数据删除管理器利用重复数据删除队列中的请求对卷中的每一个的重复数据删除值进行检查。具有较高的重复数据删除值的备用列表卷被选择以在具有较低的重复数据删除值的其它备用列表卷

之前进行重复数据删除。然而,如以下相对于图3和4更详细地解释的,优选列表上的卷在备用列表上的卷之前被选择。

[0049] 动作212包括将一个或多个卷从备用列表移动到优选列表。因此,在一个方面中,根据重复数据删除可能性和可用系统资源,一些卷随着时间而从一个列表移动到另一个列表。动作212本身可以被认为是根据重复数据删除可能性来调度重复数据删除操作的方式,因为在这个示例中,对于优选列表卷的请求在对于备用列表卷的请求之前进行处理。

[0050] 方法200核查卷的重复数据删除速率是否可以跟上该卷的实际数据改变速率。如果卷的重复数据删除可以跟上,则该卷可以是优选列表的候选。例如,如果卷的改变日志在预定的时间段内没有溢出,并且如果系统具有足够可用的资源以确保该卷(以及优选列表中的任意其它卷)可以以及时的方式进行重复数据删除,则可以将该卷从备用列表移动到优选列表。在以上示例中,当卷的改变日志达到预定容量阈值(例如,20%整)时,重复数据删除管理器将重复数据删除请求放置在重复数据删除队列中。当请求位于等待完成的重复数据删除队列中时,相应的卷的改变日志从20%增大,可能甚至到100%,其中,它被称为已溢出。溢出的改变日志在该卷的下一次重复数据删除操作时被清除。改变日志溢出通常可能是不希望的,因为当改变日志溢出时,重复数据删除管理器可能不能检测到一些重复并且不能将一些指纹添加到指纹数据库。因此,对于具有可实现的节省可能性的卷,改变日志溢出可能减小提供的节省。可以以任意适当的方式向重复数据删除管理器通知改变日志溢出,诸如通过利用当改变日志溢出时发送到重复数据删除管理器并且此后周期性地发送到重复数据删除管理器的通知。

[0051] 具有较高的重复数据删除可能性的卷可能期望被更经常地进行重复数据删除,并且在重复数据删除队列中被较早地进行,因此,不太可能使改变日志溢出。因此,在一些实施例中,自改变日志溢出以来的时间跨度可以被用作识别具有最佳节省可能性的卷的方式。

[0052] 然而,实施例的范围不限于识别要移动到优选列表的卷的任意特定技术。例如,其它实施例可以使将卷移动到优选列表的标准仅基于重复数据删除值、或者基于在预定的时间跨度内执行的重复数据删除的数量、或者任意其它适当的度量。

[0053] 动作212还可以包括将一些备用列表卷从备用列表移动到第三列表,在该第三列表中,卷不被考虑用于重复数据删除。可以将具有最小量的重复数据删除节省可能性的卷移动到第三列表。在一个示例实施例中,可以将预定的时间跨度由于改变日志溢出而留下的卷移动到第三列表,然而实施例的范围可以包括用于确定卷应被移动到第三列表的任意适当的技术。

[0054] 各种实施例可以添加、省略、重新布置或修改方法200的动作。例如,一些实施例包括每当不同的现有卷或者不同的新卷呈现给重复数据删除管理器就执行动作202、204、206、208。此外,一些实施例还可以包括在例如当系统资源减少时的情况下将卷从优选列表移动到备用列表。因此,一些实施例通过将卷从备用列表移动到优选列表并可能再次返回到备用列表来提供适应于系统资源的可变可用性的连续操作。这样的实施例可以使用任意适当的技术来选择要移动到备用列表的优选列表卷,包括例如选择具有最低的重复数据删除可能性的一个或多个优选列表卷。移回到备用列表的优选列表卷的数量可以基于例如系统资源改变到的量来确定。

[0055] 系统资源的改变可以指的是系统中的若干个不同的改变。系统资源的一种类型的改变包括将更多的盘和/或处理器添加到系统,这可以被看作为系统资源的增加,因为这样的添加允许系统进行更多处理。类似地,盘或处理器的减少可以被看作为系统资源的减少。资源的另一改变包括当用户工作负荷增加或减少时,这导致剩余系统资源的减少或增加。

[0056] 取指纹是特定于重复数据删除的,而诸如压缩的其它技术可以放弃取指纹,并取而代之的是对所有或几乎所有的添加数据进行压缩。此外,在使用压缩的示例系统中,该系统可以计算压缩的空间节省可能性,并根据所计算的空间节省可能性来分配压缩处理资源。

[0057] 图3是根据一个实施例改动的用于在众多卷之间分配用于重复数据删除的资源的示例方法300的图示。再一次地,图3集中于重复数据删除,但理解的是,图3的原理也应用于其它空间节省技术。方法300可以由存储系统中的存储服务器或其它地方处的空间节省管理器执行。

[0058] 在动作302中,重复数据删除管理器产生重复数据删除请求,并将该请求放置在重复数据删除队列中。在以上示例中,当卷的改变日志达到20%整时,重复数据删除管理器将请求放置在卷的重复数据删除队列中。然而,实施例的范围不限于该技术。各种实施例可以使用任意标准或标准的组合来确定何时产生重复数据删除请求。

[0059] 在动作304中,辨别该请求是否是针对优选列表中的卷。如果该请求是针对优选列表中的卷,则在动作306处,重复数据删除管理器在其它卷之前调度该卷上的重复数据删除。动作306可以包括将比对于备用列表上的卷的另一个请求高的优先级指派给该请求。

[0060] 如果该请求是针对不在优选列表上的卷,则在动作308处,重复数据删除管理器基于备用列表内的重复数据删除节省可能性来调度该卷的重复数据删除。换句话说,如果备用列表中的其它卷具有较高的可能性并且具有挂起(pending)的重复数据删除请求,则动作302的请求可以等待在其它请求后面进行处理。另外,该请求将在任意优选列表卷被处理之后(如果有的话)进行处理。

[0061] 注意,以上在动作308处,可以根据相应的卷的重复数据删除值来分配至少在备用列表内的重复数据删除资源。然而,各种实施例可以以比仅仅按照重复数据删除值对卷进行排列的更大的灵活性和复杂性来处理分配。在一个示例中,两个或更多个卷具有类似的重复数据删除值(例如,分别为25%和30%),并且假定它们的重复数据删除节省可能性也是类似的。因此,这样的示例实施例可以根据哪个卷具有最大的改变日志来在这两个卷之间分配重复数据删除资源。以这种方式,重复数据删除管理器的工作是最小化这些卷中的一个具有改变日志溢出的机会。为了进一步提升重复数据删除处理内的效率,可以改动其它修改和添加以供在一些实施例中使用。

[0062] 图4是根据实施例改动的在卷之间分配空间节省资源的示例方法400的图示。鉴于图3被更概念性地示出,图4意图阐明空间节省管理器遍历其中具有众多条目的重复数据删除队列的特定技术。因此,图4阐明了根据以上给出的原理从队列中选择重复数据删除请求以首先将资源分配给更有前景的卷的至少一种方式。图4的原理可以应用于实现其它空间节省技术(诸如压缩等)的系统。

[0063] 在动作402处,重复数据删除引擎访问重复数据删除请求队列。在一些实施例中,重复数据删除操作可以仅在一天的某些时间运行,和/或可以受在任意给定时间可以执行

的数量限制。因此,动作402可以在重复数据删除操作开始的某一预定时间或者事件处执行。

[0064] 在动作404处,重复数据删除引擎对于是否存在对于优选列表中的卷的重复数据删除请求来检查队列。如果存在对于优选列表卷的请求,则在动作406处,重复数据删除引擎对该卷进行重复数据删除。在一些情况下,在队列中对于优选列表卷可以存在多于一个的请求。各种实施例可以使用任意适当的技术来确定首先完成优选列表请求中的哪些。例如,重复数据删除引擎可以基于任意数量的度量来在优选列表卷之间进行选择,所述度量诸如请求被接收的次序、改变日志大小、重复数据删除值等。

[0065] 如果不存在对于优选列表卷的重复数据删除请求,则在动作408处,重复数据删除引擎对于是否存在对于备用列表中的卷的请求来检查队列。如果不存在对于备用列表卷的请求,则重复数据删除引擎可以返回到动作402,或者简单地结束重复数据删除处理,直到稍后的时间为止。

[0066] 如果确定对于备用列表卷存在至少一个重复数据删除请求,则重复数据删除引擎选择与具有最高的重复数据删除可能性的卷对应的请求。

[0067] 各种实施例可以添加、省略、重新布置或修改图4的一个或多个动作。例如,方法400可以连续地或者仅在某些时间执行,并且在任意情况下,可以在重复数据删除操作停止之前循环多次。实际上,一些系统允许在一天的任意时间期间进行重复数据删除操作,但是将任意一个时间的重复数据删除操作的数量限制为八次。因此,一些实施例可以包括在系统内并发运行的多个重复数据删除操作。当然,这仅仅是个示例,并且在各种实施例中,可以实现约束重复数据删除操作的时间和/或多重性的任意适当的技术。

[0068] 所描述的实施例可以包括优于其它技术的一个或多个优点。例如,各种实施例提供计算空间节省可能性的技术,从而将智能引入若干个卷之间的选择。此外,各种实施例有利于对具有产生大量的空间节省的较高的可能的卷执行空间节省处理。进行这样的优先是以对于空间节省显示出较小的可能性的卷为代价的。因此,这样的实施例指导资源,其中这些资源被更高效地和有效地使用。相比之下,一些常规的重复数据删除系统可能仅根据重复数据删除请求被接收的次序来执行重复数据删除,并且当资源不可用时,卷中的一些可能不进行重复数据删除。然而,不进行重复数据删除的卷中的一些可能是对于节省具有最大可能性的卷。

[0069] 各种实施例不仅对存储系统增添了效率,而且也可以增添可靠性。例如,购买空间节省应用的客户可能期望能够依赖于系统中的一定量的空间节省。然而,常规的存储系统可能提供或者可能不能提供期望的空间节省,这依赖于具有最高的重复数据删除可能性的卷实际上是否进行重复数据删除。本公开的一些实施例确保,具有较高的空间节省可能性的卷首先进行重复数据删除,从而提供更一致的空间节省。当客户使用空闲的存储空间时,客户立即受益于空间节省。客户还通过在配给决策时能够依赖于一致的节省来长期受益。在以上示例中,优选列表允许通过对一组固定的卷给予优先来实现长期益处。

[0070] 当经由计算机可执行指令实现时,本公开的实施例的各个要素本质上是定义这样的各个要素的操作的软件代码。可执行指令或软件代码可以从有形可读介质(例如,硬盘驱动器介质、光学介质、RAM、EPROM、EEPROM、磁带介质、盒介质、闪存、ROM、记忆棒、网络存储设备等)获得。实际上,可读介质可以包括可以存储信息的任意介质。网络存储系统中的存储

服务器(例如,图1的服务器102)、存储子系统(例如,子系统104和106)或其它地方处的一个或多个处理器可以执行这样的指令以提供方法200、300、400(图2-4)的动作中的一个或多个。

[0071] 前面概述了几个实施例的特征,使得本领域的技术人员可以更好地理解本公开的方面。本领域的技术人员应意识到,他们可以容易地将本公开用作设计或修改用于执行相同目的和/或实现本文中介绍的实施例的相同优点的其它处理和结构的基础。本领域的技术人员还应认识到,这样的等同构造没有偏离本公开的精神和范围,并且他们可以在不脱离本公开的精神和范围的情况下进行各种改变、替换和更改。

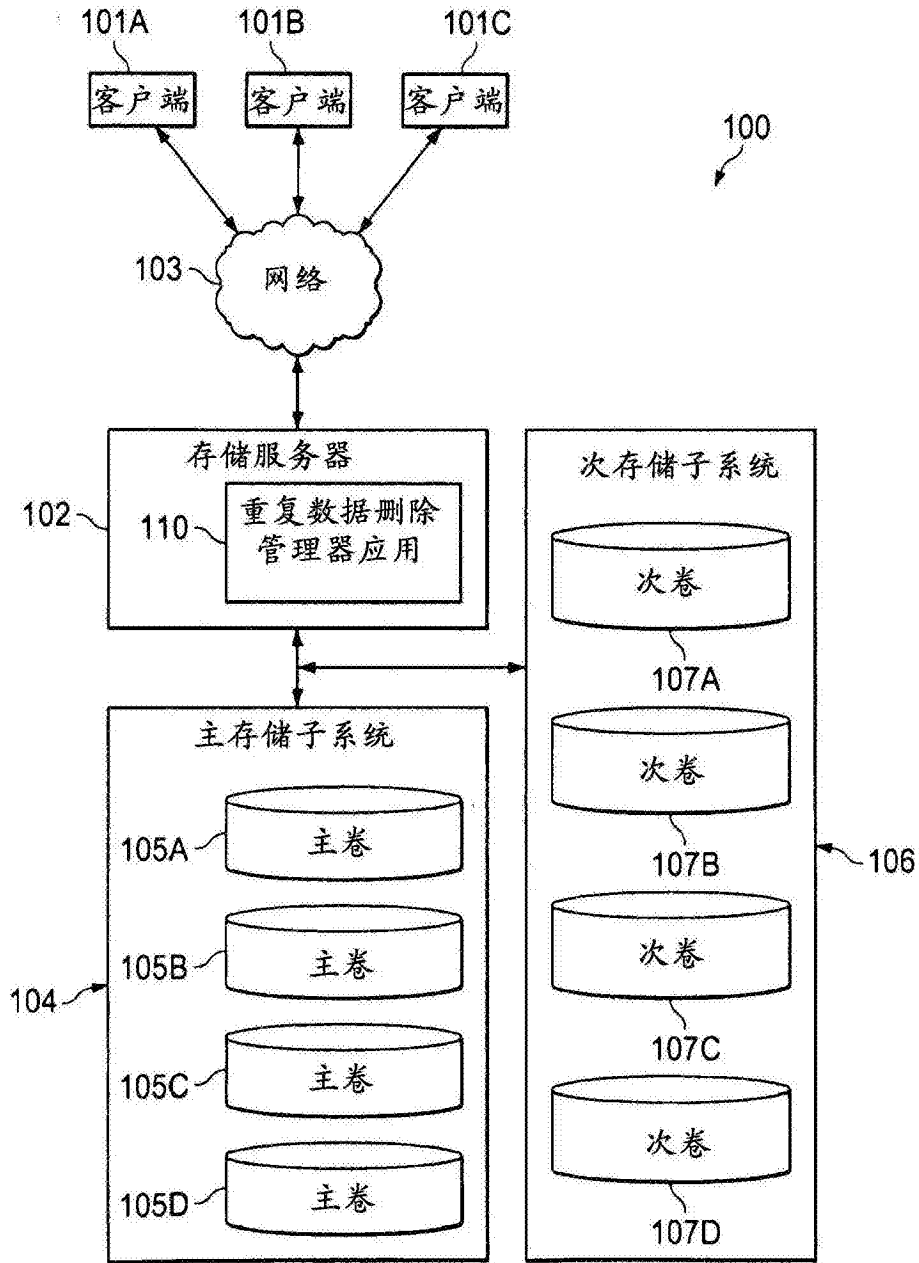


图1

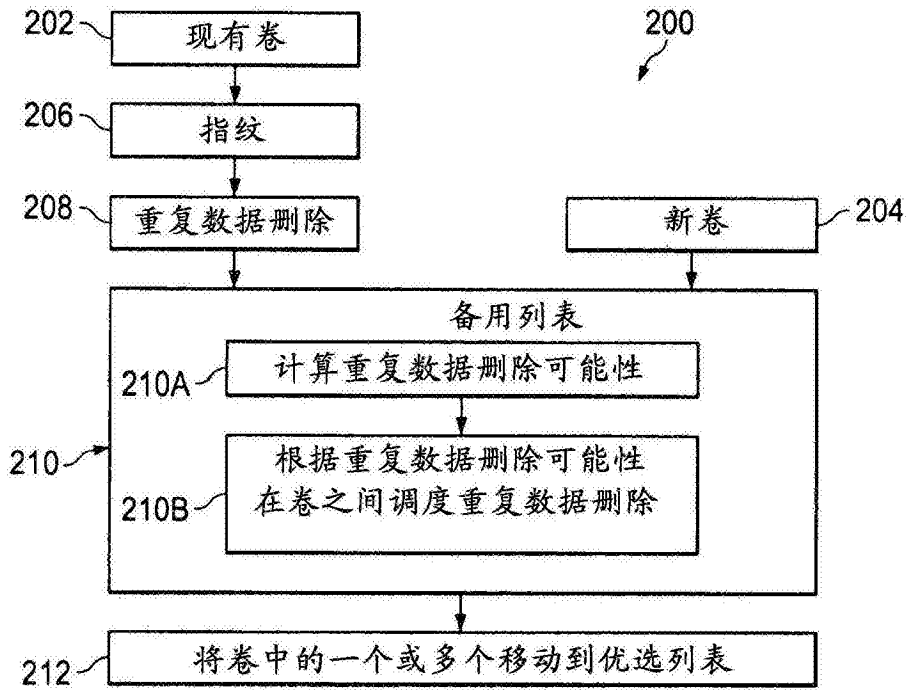


图2

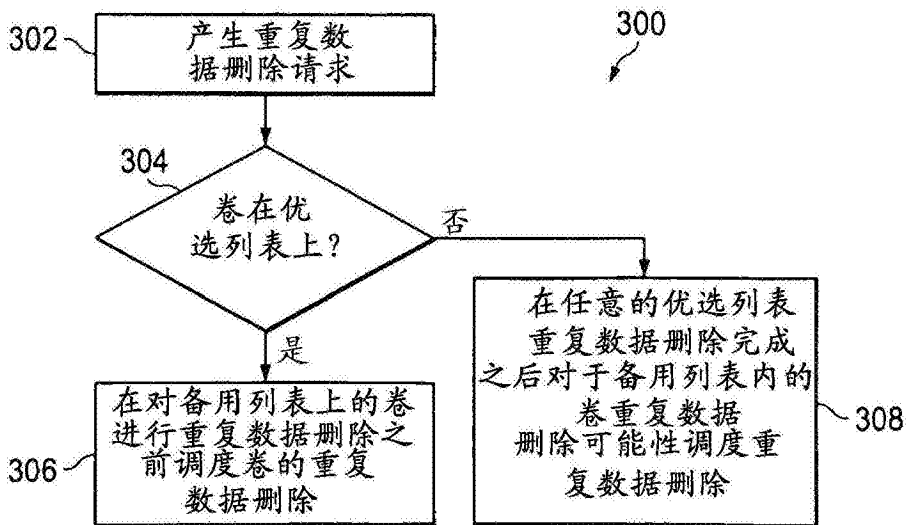


图3

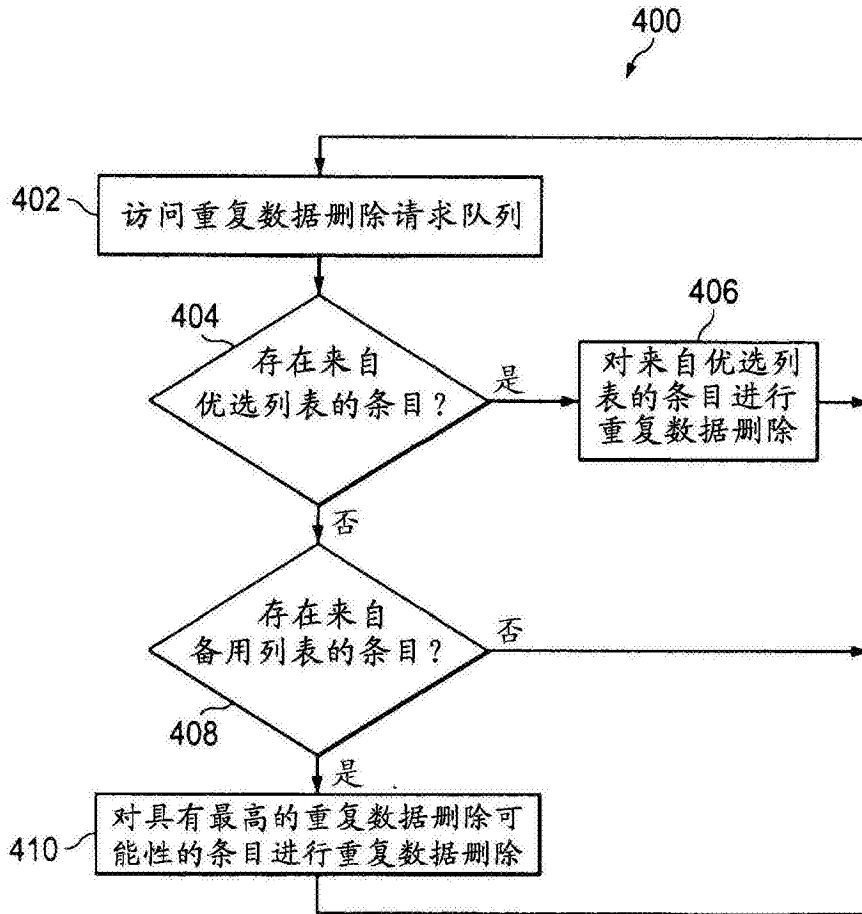


图4