



(12)发明专利申请

(10)申请公布号 CN 109144431 A

(43)申请公布日 2019.01.04

(21)申请号 201811157995.7

(22)申请日 2018.09.30

(71)申请人 华中科技大学

地址 430074 湖北省武汉市洪山区珞喻路
1037号

申请人 腾讯科技(深圳)有限公司

(72)发明人 周可 张煜 王桦 吉永光 程彬

(74)专利代理机构 北京三高永信知识产权代理
有限责任公司 11138

代理人 祝亚男

(51)Int.Cl.

G06F 3/06(2006.01)

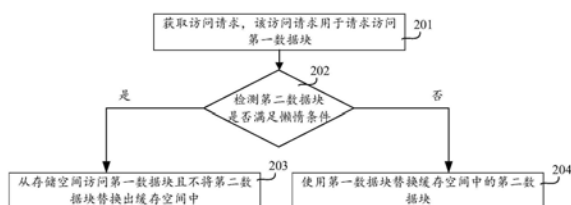
权利要求书3页 说明书13页 附图7页

(54)发明名称

数据块的缓存方法、装置、设备及存储介质

(57)摘要

本申请公开了一种数据块的缓存方法、装置、计算机设备及存储介质,属于存储技术领域。所述方法包括:获取访问请求,所述访问请求用于请求访问第一数据块;当所述第一数据块在所述缓存空间中未命中时,检测第二数据块是否满足懒惰条件,所述第二数据块是所述缓存空间中的备选剔除块;当所述第二数据块满足所述懒惰条件时,从所述存储空间访问所述第一数据块且不将所述第二数据块替换出所述缓存空间中。本申请能够极大地延长第二数据块在缓存空间中的停留时长,提高有用数据块的命中率,从而提升存储系统的读写性能。



1. 一种数据块的缓存方法,用于存储系统中,所述存储系统包括存储空间和缓存空间,其特征在于,所述方法包括:

获取访问请求,所述访问请求用于请求访问第一数据块;

当所述第一数据块在所述缓存空间中未命中时,检测第二数据块是否满足懒惰条件,所述第二数据块是所述缓存空间中的备选剔除块,所述懒惰条件是根据重新访问概率确定将所述第二数据块延迟替换出所述缓存空间的条件;

当所述第二数据块满足所述懒惰条件时,从所述存储空间访问所述第一数据块且不将所述第二数据块替换出所述缓存空间中。

2. 根据权利要求1所述的方法,其特征在于,所述存储系统存储有数据块标识列表BIL,所述BIL存储有历史访问的n个数据块的块元数据,n为正整数;

所述检测所述第二数据块是否满足懒惰条件,包括:

当所述第一数据块不存在于所述BIL中时,检测所述第二数据块是否满足第一懒惰条件;

当所述第一数据块存在于所述BIL中时,检测所述第二数据块是否满足第二懒惰条件;

其中,所述第一懒惰条件所指示的重新访问概率小于所述第二懒惰条件所指示的重新访问概率。

3. 根据权利要求2所述的方法,其特征在于,

所述第一懒惰条件包括:所述第二数据块的所述标记值大于第一阈值,所述标记值是用于衡量所述第二数据块的访问重要程度的变量;

所述第二懒惰条件包括:所述第二数据块的所述标记值大于第二阈值,且所述第二数据块在所述缓存空间中的停留时长大于K倍的块平均重用距离, $K > 0$ 。

4. 根据权利要求3所述的方法,其特征在于,所述方法还包括;

当所述第二数据块在所述缓存空间中命中时,按照第一幅度增加所述标记值;

当所述第二数据块在所述缓存空间中因满足所述懒惰条件未被替换时,按照第二幅度减少所述标记值。

5. 根据权利要求2至4任一所述的方法,其特征在于,所述方法还包括:

当所述第二数据块满足所述第一懒惰条件时,将所述第一数据块的块元数据添加至所述BIL的列表头部,并将所述BIL的列表尾部的第三数据块的块元数据进行删除;

当所述第二数据块满足所述第二懒惰条件时,将所述第一数据块的块元数据移动至所述BIL的列表头部。

6. 根据权利要求2至4任一所述的方法,其特征在于,所述方法还包括:

当所述第二数据块不满足所述第一懒惰条件或所述第二懒惰条件时,使用所述第一数据块替换所述第二数据块。

7. 根据权利要求2至4任一所述的方法,其特征在于,所述存储系统存储有懒惰替换列表LEL,所述LEL存储有位于所述缓存空间中的n个数据块的块元数据,n为正整数;所述方法还包括:

当所述第二数据块不满足所述第一懒惰条件时,将所述第二数据块的块元数据从所述LEL的列表尾部删除;

当所述第二数据块不满足所述第二懒惰条件时,将所述第二数据块的块元数据从所述

LEL的列表尾部删除,且将所述第二数据块的块元数据移动至所述BIL的列表头部。

8.根据权利要求1至4任一所述的方法,其特征在于,所述存储系统存储有懒惰替换列表LEL,所述LEL存储有位于所述缓存空间中的n个数据块的块元数据,n为正整数;所述方法还包括:

当所述第二数据块不满足所述懒惰条件时,将所述第一数据块的块元数据移动至所述LEL的列表头部。

9.一种数据块的缓存装置,其特征在于,应用于存储系统中,所述存储系统包括存储空间和缓存空间,所述装置包括:

获取模块,用于获取访问请求,所述访问请求用于请求访问第一数据块;

检测模块,用于当所述第一数据块在所述缓存空间中未命中时,检测第二数据块是否满足懒惰条件,所述第二数据块是所述缓存空间中的备选剔除块,所述懒惰条件是根据重新访问概率确定将所述第二数据块延迟替换出所述缓存空间的条件;

访问模块,用于当所述第二数据块满足所述懒惰条件时,从所述存储空间访问所述第一数据块且不将所述第二数据块替换出所述缓存空间中。

10.根据权利要求9所述的装置,其特征在于,所述存储系统存储有数据块标识列表BIL,所述BIL存储有历史访问的n个数据块的块元数据,n为正整数;

所述检测模块,用于当所述第一数据块不存在于所述BIL中时,检测所述第二数据块是否满足第一懒惰条件;当所述第一数据块存在于所述BIL中时,检测所述第二数据块是否满足第二懒惰条件;

其中,所述第一懒惰条件所指示的重新访问概率小于所述第二懒惰条件所指示的重新访问概率。

11.根据权利要求10所述的装置,其特征在于,

所述第一懒惰条件包括:所述第二数据块的所述标记值大于第一阈值,所述标记值是用于衡量所述第二数据块的访问重要程度的变量;

所述第二懒惰条件包括:所述第二数据块的所述标记值大于第二阈值,且所述第二数据块在所述缓存空间中的停留时长大于K倍的块平均重用距离。

12.根据权利要求10至11任一所述的装置,其特征在于,所述装置还包括:

元数据修改模块,用于当所述第二数据块满足所述第一懒惰条件时,将所述第一数据块的块元数据添加至所述BIL的列表头部,并将所述BIL的列表尾部的第三数据块的块元数据进行删除;以及当所述第二数据块满足所述第二懒惰条件时,将所述第一数据块的块元数据移动至所述BIL的列表头部。

13.根据权利要求10至11任一所述的装置,其特征在于,所述存储系统存储有懒惰替换列表LEL,所述LEL存储有位于所述缓存空间中的n个数据块的块元数据,n为正整数;所述装置还包括:

元数据修改模块,用于当所述第二数据块不满足所述第一懒惰条件时,将所述第二数据块的块元数据从所述LEL的列表尾部删除;以及当所述第二数据块不满足所述第二懒惰条件时,将所述第二数据块的块元数据从所述LEL的列表尾部删除,且将所述第二数据块的块元数据移动至所述BIL的列表头部。

14.一种应用于存储系统中的计算机设备,其特征在于,所述存储系统包括存储空间和

缓存空间,所述计算机设备包括:

处理器;

与所述处理器相连的收发器;

用于存储处理器可执行指令的存储器;

其中,所述处理器被配置为加载并执行所述可执行指令以实现如上权利要求1至8任一所述的数据块的缓存方法。

15.一种计算机可读存储介质,其特征在于,所述可读存储介质中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现如权利要求1至8任一所述的数据块的缓存方法。

数据块的缓存方法、装置、设备及存储介质

技术领域

[0001] 本申请涉及存储技术领域,特别涉及一种数据块的缓存方法、装置、设备及存储介质。

背景技术

[0002] 数据存储系统包括:存储空间和缓存空间。存储空间存储有全量数据块,缓存空间用于存储被频繁访问的热度数据块。热度数据块是全量数据块的子集。

[0003] 在数据存储系统中采用缓存算法来移入或移出缓存空间中的热度数据块。在缓存空间中已经存储满的场景下,对于缓存空间中每个未命中的数据块,缓存算法将该数据块从存储空间中读取后,作为新的热度数据块与缓存空间中的一个旧热度数据块进行交换。

[0004] 由于典型的使用场景下,存在很多热度数据块的重复访问在很长时间后才能发生。如果缓存算法只是简单的在每次未命中时替换缓存空间中的热度数据块,新存入的热度数据块会污染缓存空间,而这个热度数据块在将来的缓存过程中可能永远不会被命中,从而降低缓存空间的使用效率。

发明内容

[0005] 本申请实施例提供了一种数据块的缓存方法、装置、设备及存储介质,可以解决在缓存空间中每个未命中的数据块均进行替换时,新存入的热度数据块会污染缓存空间,而这个热度数据块在将来的缓存过程中可能永远不会被命中,从而降低缓存空间的使用效率的问题。所述技术方案如下:

[0006] 根据本申请的一个方面,提供了一种数据块的缓存方法,用于存储系统中,所述存储系统包括存储空间和缓存空间,所述方法包括:

[0007] 获取访问请求,所述访问请求用于请求访问第一数据块;

[0008] 当所述第一数据块未在所述缓存空间中命中时,检测第二数据块是否满足懒惰条件,所述第二数据块是所述缓存空间中的备选剔除块,所述懒惰条件是用于衡量所述第一数据块的重新访问概率的条件;

[0009] 当所述第二数据块满足所述懒惰条件时,从所述存储空间读取所述第一数据块且不将所述第一数据块添加至所述缓存空间中;

[0010] 当所述第二数据块不满足所述懒惰条件时,使用所述第二数据块替换所述缓存空间中的第二数据块。

[0011] 根据本申请的另一方面,提供了一种数据块的缓存装置,应用于存储系统中,所述存储系统包括存储空间和缓存空间,所述装置包括:

[0012] 获取模块,用于获取访问请求,所述访问请求用于请求访问第一数据块;

[0013] 检测模块,用于当所述第一数据块未在所述缓存空间中命中时,检测所述第一数据块是否满足懒惰条件,所述懒惰条件是用于衡量所述第一数据块的重新访问概率的条件;

[0014] 读取模块,用于当所述第一数据块满足所述懒惰条件时,从所述存储空间读取所述第一数据块且不将所述第一数据块添加至所述缓存空间中。

[0015] 根据本申请的另一方面,提供了一种应用于存储系统中的计算机设备,所述存储系统包括存储空间和缓存空间,所述计算机设备包括:

[0016] 处理器;

[0017] 与所述处理器相连的收发器;

[0018] 用于存储处理器可执行指令的存储器;

[0019] 其中,所述处理器被配置为加载并执行所述可执行指令以实现如上所述的数据块的缓存方法。

[0020] 根据本申请的另一方面,提供了一种计算机可读存储介质,所述计算机可读存储介质中存储有至少一条指令、至少一段程序、代码集或指令集,所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现如上所述的数据块的缓存方法。

[0021] 本申请实施例提供的技术方案带来的有益效果至少包括:

[0022] 通过当第一数据块在缓存空间中未命中时,检测缓存空间中的第二数据块是否满足懒惰条件,当第二数据块满足懒惰条件时,从存储空间读取第一数据块且暂不将第二数据块替换出缓存空间中,能够极大地延长第二数据块在缓存空间中的停留时长,提高有用数据块的命中率,从而提升存储系统的读写性能。

附图说明

[0023] 为了更清楚地说明本申请实施例中的技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0024] 图1是本申请一个示例性实施例提供的存储系统的结构示意图;

[0025] 图2是本申请一个示例性实施例提供的数据块的缓存方法的流程图;

[0026] 图3是本申请另一个示例性实施例提供的数据块的缓存方法的流程图;

[0027] 图4是本申请另一个示例性实施例提供的数据块的缓存方法的流程图;

[0028] 图5是本申请一个示例性实施例提供的索引表、元数据区和数据区的结构示意图;

[0029] 图6是本申请一个示例性实施例提供的第一懒惰条件的示意图;

[0030] 图7是本申请一个示例性实施例提供的第二懒惰条件的示意图;

[0031] 图8是本申请一个示例性实施例提供的数据块的缓存方法的实施示意图;

[0032] 图9是本申请一个示例性实施例提供的数据块的缓存装置的结构框图;

[0033] 图10是本申请一个示例性实施例提供的数据块的缓存装置的结构框图;

[0034] 图11是本申请另一个示例性实施例提供的计算机设备的结构框图。

具体实施方式

[0035] 为使本申请的目的、技术方案和优点更加清楚,下面将结合附图对本申请实施方式作进一步地详细描述。

[0036] 首先对本申请涉及的若干个术语进行介绍:

[0037] 懒惰替换:是指第一数据块在缓存空间中未命中时,条件性的将该第一数据块与缓存空间中的第二数据块(也称备选剔除块)进行替换的方式。这里的“懒惰”是相对于相关技术中,在缓存空间中每次未命中数据块时,均强制对缓存空间中的备选剔除块进行强制替换而言的。

[0038] 重用距离:是对同一个数据块的两次连续访问之间的唯一数据块的计数,或,对同一个数据块的两次连续访问之间的数据块的计数。例如,历史访问的数据块序列包括:1-2-4-5-4-3-3-2,当采用唯一数据块的计数方式时,数据块2的重用距离是3,因为在数据块的两次连续访问之间的唯一数据块集合是{3,4,5};当采用数据块的计数方式时,数据块的重用距离是5。连续访问也称重新访问,是指同一个数据块再次被访问的事件。

[0039] 在存储系统中,数据块的重用距离对缓存命中率有很大的影响。通常,不同的存储系统具有不同的重用距离特性。例如,在传统的片上高速缓存环境中,平均重用距离可能远远短于云块存储系统。

[0040] 索引表:是一种用于加速数据查找的数据结构,常见的索引表结构包括:B树、B+树、哈希散列表以及红黑树等。

[0041] 命中:当一个数据块被访问时,若该数据块存在于缓存空间中,即认为在缓存空间中命中了该数据块。

[0042] 图1示出了本申请一个示例性实施例提供的存储系统100的结构框图。该存储系统100包括:控制设备120、缓存空间140和存储空间160。

[0043] 控制设备120是用于控制缓存空间140对被频繁访问的数据块142进行存储的计算机设备。可选地,控制设备120中运行有缓存算法,该缓存算法用于实现本申请各个实施例中提供的数据块的缓存方法。

[0044] 缓存空间140是用于对存储空间160中被频繁访问的数据块142进行缓存的存储逻辑单元。缓存空间140的读写性能优于存储空间160的读写性能。缓存空间140可采用非易失性缓存设备来构建。非易失性缓存设备包括:固态硬盘(Solid State Drives,SSD)、相变存储(Phase-Change Memory,PCM)、自旋转移扭矩磁随机存取器(Shared Transistor Technology Magnetic Random Access Memory,STT-MRAM)及电阻式内存(ReRAM)等。

[0045] 存储空间160是用于对全量的数据块164进行存储的存储逻辑单元。存储空间160可采用非易失性存储设备来构建。常见的非易失性存储设备为机械硬盘。在存储空间160中存在多个存储卷162,每个存储卷162中存储有多个数据块164。可选地,数据块的大小是相同或不同的,比如数据块的大小可以是4KB、8KB、16KB、32KB等。每个存储卷162对应有各自的卷ID,每个数据块164对应有各自的块ID。

[0046] 在一些可选的实施例中,存储系统100还包括接入设备110,该接入设备110用于接收各个终端180发送的数据块的访问请求,并将各个访问请求按照接收时间由早到晚的顺序存储在请求列表112中。

[0047] 示意性,各个客户端180可以向接入设备110发送数据块的访问请求,该数据块的访问请求可以是读请求或写请求。控制设备120根据该访问请求对缓存空间140和存储空间160中的数据块进行访问。当缓存空间140中命中该数据块时,控制设备120优先从缓存空间140访问该数据块;当缓存空间140中未命中该数据块时,控制设备120从存储空间160访问

该数据块。

[0048] 在相关技术中,若客户端向存储系统请求访问数据块,而该数据块在缓存空间中未命中时,缓存算法会将每个未命中的数据块从存储空间移入缓存空间中(若缓存空间已经满,则需要将相应的备选剔除块移出缓存空间)。但是这种简单地将每个未命中的数据块替换至缓存空间中的方式,会使得一些重新访问可能性很低的数据块进入缓存空间,从而污染缓存空间,但将来可能永远都不会被命中。

[0049] 另外,由于新型的非易失性缓存介质都具有写次数限制的特性,当写次数达到一定次数后即会寿命耗尽,因此当缓存空间采用这种新型的非易失性缓存介质构建时,上述缓存算法也会很快地耗尽缓存介质的寿命。

[0050] 本申请实施例希望提供一种尽可能将有价值的数据块保留在缓存空间的技术方案。本申请实施例提供的数据块的缓存方法,可以应用于具有缓存空间和存储空间的存储系统中。典型的,该存储系统是云块存储系统。云块存储系统为用户提供一种弹性的块级别云存储服务,云块存储系统将同时服务于拥有多种工作负载的数万个租户。

[0051] 图2示出了本申请一个示例性实施例提供的数据块的缓存方法的流程图。本实施例以该方法应用于图1所示的存储系统中来举例说明。该方法包括:

[0052] 在步骤201中,获取访问请求,该访问请求用于请求访问第一数据块;

[0053] 控制设备从请求列表中读取一个访问请求,请求列表中按序存储有多个访问请求。示例性的,控制设备从请求列表中读取接收时间最早的访问请求,该访问请求用于请求访问第一数据块。

[0054] 示意地,该访问请求中携带有第一数据块的标识信息,该标识信息包括:卷ID和块ID,或者,块ID。控制设备根据标识信息检测第一数据块是否存在于缓存空间中。

[0055] 当第一数据块在缓存空间中命中时,控制设备在缓存空间中访问该第一数据块。当第一数据块在缓存空间中未命中时,执行步骤202。

[0056] 在步骤202中,当第一数据块在缓存空间中未命中时,检测第二数据块是否满足懒惰条件,第二数据块是缓存空间中的备选剔除块,懒惰条件是根据重新访问概率确定将第二数据块延迟替换出缓存空间的条件;

[0057] 示意性的,当缓存空间中的各个数据块采用链表按照重新访问概率进行排序存储时,第二数据块是位于链表尾部的数据块。

[0058] 懒惰条件又称为懒惰替换条件。可选地,懒惰条件是根据第二数据块的重新访问概率,决策是否要将第二数据块延迟替换出缓存空间中的条件;可选地,懒惰条件是当第二数据块的重新访问概率高于期望阈值时,确定将第二数据块延迟替换出缓存空间中的条件;可选地,懒惰条件是根据第二数据块的重新访问分值和/或重用距离来确定第二数据块的重新访问概率,根据第二数据块的重新访问概率决策是否要将第二数据块延迟替换出缓存空间中的条件。

[0059] 在步骤203中,当第二数据块满足懒惰条件时,从存储空间访问第一数据块且不将第二数据块替换出缓存空间中;

[0060] 当第二数据块满足懒惰条件时,代表第二数据块的重新访问概率较高,控制设备从存储空间中读取第一数据块,暂不将第二数据块替换出缓存空间中。

[0061] 在步骤204中,当第二数据块不满足懒惰条件时,使用第一数据块替换缓存空间中

的第二数据块。

[0062] 当第二数据块不满足懒惰条件时,代表第二数据块的重新访问概率较低,控制设备从存储空间读取第一数据块后,使用第一数据块替换缓存空间中的第二数据块。

[0063] 综上所述,本实施例提供的方法,通过当第一数据块在缓存空间中未命中时,检测第二数据块是否满足懒惰条件,当第二数据块满足懒惰条件时,从存储空间读取第一数据块且暂不将第二数据块替换出缓存空间中,能够极大地延长第二数据块在缓存空间中的停留时长,提高有用数据块的命中率,从而提升存储系统的读写性能。

[0064] 在基于图2的可选实施例中,当缓存空间中未命中第一数据块时,如果缓存空间的第二数据块满足懒惰条件,则暂不将第二数据块替换出缓存空间,从而延长第二数据块在缓存空间中的停留时长。对于未替换入缓存空间的第一数据块,使用数据块标识列表(Block Identifier List,BIL)记录该第一数据块的历史访问情况。对于存储在BIL中的第一数据块,在下次被访问时具有更高的概率被替换入缓存空间中。通过该方式,能够更加高效的筛选出对缓存空间更有益的数据块,同时也可以避免对缓存设备不必要的写入。

[0065] 图3示出了本申请一个示例性实施例提供的数据块的缓存方法的流程图。本实施例以该方法由图1中的控制设备120执行来举例说明。该方法包括:

[0066] 在步骤301中,获取访问请求,该访问请求用于请求访问第一数据块;

[0067] 控制设备从请求列表中读取一个访问请求,请求列表中按序存储有多个访问请求。示例性的,控制设备从请求列表中读取接收时间最早的第*i*个访问请求,该第*i*个访问请求用于请求访问第一数据块。

[0068] 在步骤302中,检测第一数据块是否存在于缓存空间中;

[0069] 示意地,该访问请求中携带有第一数据块的标识信息,该标识信息包括:卷ID和块ID,或者,块ID。控制设备根据标识信息检测第一数据块是否存在于缓存空间中。当存在于缓存空间时,即为在缓存空间中命中该第一数据块。

[0070] 当第一数据块在缓存空间中命中时,控制设备在缓存空间中访问该第一数据块。当第一数据块在缓存空间中未命中时,执行步骤303。

[0071] 在步骤303中,当第一数据块在缓存空间中未命中时,检测第一数据块是否存在于BIL中;

[0072] BIL存储有历史访问的*n*个数据块的块元数据。可选地,块元数据包括但不限于:数据块的块ID、数据块所在存储卷的卷ID、数据块的重用距离、数据块的标记(flag)值、数据块在缓存空间中的停留时长中的至少一项。

[0073] 其中,flag值是用于表示数据块在重新访问维度的重要性的分值,也称重新访问分值,由控制设备根据数据块的历史重新访问数据动态维护。停留时长是按照数据块在缓存空间中存储后,经历的数据块访问次数来计量的。比如一个数据块2在缓存空间中存储后,后续的数据块访问序列为:1-2-3-6-7-5-4-7,则该数据块的停留时长为8。

[0074] 控制设备根据第一数据块的块ID,查询第一数据块是否存储于BIL中。当第一数据块不存在于BIL中时,进入步骤304;当第一数据块存在于BIL中时,进入步骤305。

[0075] 本实施例中提供有两个懒惰条件:第一懒惰条件和第二懒惰条件,第一懒惰条件所指示的重新访问概率小于第二懒惰条件所指示的重新访问概率。也即,第二懒惰条件比第一懒惰条件具有更强的限定,第二数据块需要具有更大的重新访问概率,才能继续停留

在缓存空间中。

[0076] 在步骤304中,当第一数据块不存在于BIL中时,检测第二数据块是否满足第一懒惰条件;

[0077] 示意性的,第一懒惰条件包括:

[0078] 1、第二数据块的flag值大于预设值(比如0)。

[0079] 当第二数据块满足第一懒惰条件时,进入步骤306;当第二数据块不满足第一懒惰条件时,进入步骤307。

[0080] 在步骤305中,当第一数据块存在于BIL中时,检测第二数据块是否满足第二懒惰条件;

[0081] 示意性的,第二懒惰条件包括:

[0082] 1、第二数据块的flag值大于预设值(比如0);且,

[0083] 2、第二数据块在缓存空间中的停留时长大于K倍的数据块平均重用距离。可选地,K为可变参数,随存储系统的负载大小变化而变化。

[0084] 在步骤306中,从存储空间访问第一数据块且不将第二数据块替换出缓存空间中;

[0085] 控制设备从存储空间访问第一数据块,对第一数据块进行读取或写入。但暂不将第一数据块存入缓存空间,也不将第二数据块替换出缓存空间中。将第二数据块继续保留在缓存空间中进行缓存。

[0086] 在步骤307中,使用第一数据块替换缓存空间中的第二数据块。

[0087] 控制设备从存储空间访问第一数据块,对第一数据块进行读取或写入。同时将第二数据块移出缓存空间,并将第一数据块存入缓存空间中进行缓存。

[0088] 然后,控制设备更新相应数据块的块元数据。

[0089] 综上所述,本实施例提供的方法,通过提供两个懒惰条件来进行数据块的重新访问可能性判断,当第一数据块存在于BIL时,第一数据块具有更高的概率被换入缓存空间中,从而使得更有价值的数据块存入缓存空间中,减少缓存空间中的数据块被替换出的可能性,从而提高缓存空间的利用率,也减少对缓存空间的写入次数。

[0090] 在基于图3的可选实施例中,存储系统中提供有两个列表:懒惰替换列表(Lazy Eviction List,LEL)和BIL。

[0091] LEL用于存储位于缓存空间的数据块的实体以及块元数据,数据块的实体也即数据块的数据本身。块元数据是用于描述数据块的元数据,包括但不限于:数据块的块ID、数据块所在存储卷的卷ID、数据块的重用距离、数据块的flag值、数据块在缓存空间中的停留时长中的至少一项。其中:

[0092] 块ID用于在整个存储系统或所在存储卷中唯一标识数据块。块ID可以采用数字、字母和字符中的至少一种来表示。

[0093] 卷ID用于在整个存储系统中唯一标识存储卷。

[0094] 重用距离是对同一个数据块的两次连续访问之间的唯一数据块(或数据块)的计数。

[0095] flag值是用于表示数据块在重新访问维度的重要性的分值,也称重新访问分值,由控制设备根据数据块的历史重新访问数据动态维护。在一个可能的实施例中,当一个数据块在缓存空间中被命中时,按照第一幅度增加flag值;当一个数据块在缓存空间中作为

备选剔除块且因满足懒惰条件未被替换时,按照第二幅度减少flag值。

[0096] 可选地,第一幅度小于第二幅度,比如,flag的初始值为0,当数据块在缓存空间中命中时,flag值加1;当数据块作为备选剔除块,却由于满足懒惰条件而未被替换出缓存空间时,flag值将被减半。

[0097] 停留时长是按照数据块在缓存空间中存储后,经历的数据块访问次数来计量的。比如一个数据块2在缓存空间中存储后,后续的数据块访问序列为:1-2-3-6-7-5-4-7,则该数据块的停留时长为8。

[0098] BIL用于存储历史访问过的n个数据块。BIL是一个影子链表,BIL可以仅存储数据块的卷ID以及块ID等块元信息,不需要存储数据块的实体。

[0099] 可选地,LEL和BIL具有相同的列表长度。LEL和BIL的两条链表都包含列表头部和列表尾部两个端点,数据块总是从列表头部插入链表中,从列表尾部删除。

[0100] 图4是本申请另一个示意性实施例提供的数据块的缓存方法的流程图。本实施例以该方法应用于图1所示的存储系统的控制设备中来举例说明。该方法包括:

[0101] 步骤401,从请求列表中读取一个访问请求,该访问请求用于请求访问第一数据块;

[0102] 控制设备从请求列表中读取一个访问请求,请求列表中按序存储有多个访问请求。示例性的,控制设备从请求列表中读取一个访问请求,比如接收时间最早的访问请求,该访问请求用于请求访问第一数据块。

[0103] 示意地,该访问请求中携带有第一数据块的标识信息,该标识信息包括:卷ID和块ID,或者,块ID。本实施例以卷ID和块ID来唯一标识数据块进行举例说明。

[0104] 步骤402,在索引表中查询第一数据块;

[0105] 索引表中存储有每个数据块的卷ID、块ID、以及数据块的链指针。该数据块的链指针可以指向LEL,还可以指向BIL。

[0106] 示意性的参考图5,缓存空间中存储有索引表50、元数据区51和数据区52。索引表50存储有每个数据块的卷ID、块ID和链指针。当数据块存在于LEL链表中时,链指针指向LEL链表中的某一个链表元素;当数据块存在于BIL链表中时,链指针指向BIL链表中的某一个链表元素。链表是一种数据结构,移入链表的数据从链表头部进入链表,移出链表的数据从链表尾部移出链表。在其它实施例中,LEL和BIL也可采用队列等数据结构,本实施例对此不加以限定。

[0107] 比如,卷m中的块n位于LEL链表中的链表头部,也即存在于LEL中;卷8中的块4位于BIL链表中的链表尾部,也即存在于LEL中。

[0108] LEL链表中存储有每个数据块的块元数据以及每个数据块在数据区52中的指针(可认为是数据块的实体)。BIL链表中存储有每个数据块的块元数据。

[0109] 每个数据块的数据内容存储在数据区52中。

[0110] 控制设备根据第一数据块的卷ID和块ID,在索引表中查询第一数据块的存储信息。

[0111] 步骤403,判断第一数据块是否存在于LEL;

[0112] 控制设备判断第一数据块的链指针是否指向LEL链表中的某个链表元素。

[0113] 当第一数据块存在于LEL时,认为在缓存空间中命中第一数据块,进入步骤404;当

第一数据块不存在于LEL时,认为在缓存空间中未命中第二数据块,进入步骤406。

[0114] 步骤404,从LEL访问第一数据块;

[0115] 由于缓存空间中存储有第一数据块,控制设备从LEL链表中获取第一数据块在数据区中的存储位置,根据该存储位置对第一数据块进行读取操作或写入操作。

[0116] 步骤405,将第一数据块的块元数据移动至LEL的列表头部;

[0117] 控制设备在缓存空间中命中第一数据块时,还会将第一数据块的块元数据移动至LEL的列表头部,降低第一数据块被替换出缓存空间的可能性。

[0118] 步骤406,判断第一数据块是否存在于BIL;

[0119] 当在缓存空间中未命中第一数据块时,控制设备判断第一数据块的链指针是否指向BIL链表中的某个链表元素。

[0120] 当第一数据块未存在于BIL时,表示第一数据块是最近一段时间内的首次访问数据块,采用相对宽松的第一懒惰条件对位于LEL的列表尾部的第二数据块进行懒惰替换评估,进入步骤407。

[0121] 当第一数据块存在于BIL时,表示第一数据块是最近一段时间内历史访问过的数据块,采用更为严苛的第二懒惰条件对位于LEL的列表尾部的第二数据块进行懒惰替换评估,以便增大第一数据块被替换入缓存空间的可能性,进入步骤408。

[0122] 步骤407,判断第二数据块是否满足第一懒惰条件;

[0123] 示意性的,第一懒惰条件包括:第二数据块的flag值大于0。

[0124] 当第二数据块满足第一懒惰条件时,进入步骤409;当第二数据块不满足第一懒惰条件时,进入步骤412。

[0125] 步骤408,判断第二数据块是否满足第二懒惰条件;

[0126] 示意性的,第二懒惰条件包括:第二数据块的flag值大于0,且第二数据块在LEL中的停留时长大于K倍的块平均重用距离,K为可变参数,K可以随着存储系统的负载变化而变化。比如,当负载较重时,增大K来减少替换次数;当负载较轻时,减少K来增加替换次数。

[0127] 当第二数据块满足第二懒惰条件时,进入步骤415;当第二数据块不满足第二懒惰条件时,进入步骤417。

[0128] 步骤409,从存储空间访问第一数据块,暂不将第二数据块替换出缓存空间;

[0129] 当第二数据块满足第一懒惰条件时,控制设备从存储空间读取或写入第一数据块,暂不将第一数据块插入LEL中,也即暂不采用第一数据块将第二数据块替换出缓存空间。

[0130] 步骤410,将第一数据块的块元数据移动至BIL的列表头部;

[0131] 控制设备还将第一数据块的块元数据插入至BIL的列表头部,以便在BIL中记录本次对第一数据块的访问。

[0132] 步骤411,将位于BIL的列表尾部的第三数据块的块元数据删除;

[0133] 控制设备还将第三数据块的块元数据从BIL的列表尾部删除,以保证BIL中的链表元素的总数为n个。

[0134] 步骤412,使用第一数据块替换第二数据块;

[0135] 当第二数据块不满足第一懒惰条件时,控制设备在LEL中使用第一数据块替换第二数据块,也即将第一数据块插入LEL,且将第二数据块从LEL中删除。

- [0136] 步骤413,将第一数据块的块元数据移动至LEL的列表头部;
- [0137] 在使用第一数据块替换第二数据块后,控制设备将第一数据块的块元数据插入LEL的列表头部。
- [0138] 步骤414,将位于LEL的列表尾部的第二数据块的块元数据删除;
- [0139] 控制设备还将第二数据块的块元数据从LEL的列表尾部删除,以保证LEL中的链表元素的总数为 n 个。
- [0140] 参考图6,当第一数据块是首次访问块 X 时,若第二数据块满足第一懒惰条件,则将第一数据块的块元数据插入BIL的列表头部;若第二数据块不满足第一懒惰条件,则将第一数据块的块元数据插入LEL的列表头部。
- [0141] 步骤415,从存储空间访问第一数据块,暂不将第二数据块替换出缓存空间;
- [0142] 当第二数据块满足第二懒惰条件时,控制设备从存储空间读取或写入第一数据块,暂不将第一数据块插入LEL中,也即暂不采用第一数据块将第二数据块替换出缓存空间。
- [0143] 步骤416,将第一数据块的块元数据移动至BIL的列表头部;
- [0144] 由于第一数据块的块元数据已经存在于BIL中,因此控制设备将第一数据块的块元数据移动至BIL的列表头部。
- [0145] 步骤417,使用第一数据块替换第二数据块;
- [0146] 当第二数据块不满足第二懒惰条件时,控制设备在LEL中使用第一数据块替换第二数据块,也即将第一数据块插入LEL,且将第二数据块从LEL中删除。
- [0147] 步骤418,将第一数据块的块元数据移动至LEL的列表头部;
- [0148] 在使用第一数据块替换第二数据块后,控制设备将第一数据块的块元数据插入LEL的列表头部。
- [0149] 步骤419,将位于LEL列表尾部的第二数据块的块元数据移动至BIL的列表头部;
- [0150] 控制设备还将第二数据块的块元数据从LEL的列表尾部删除,以保证LEL中的链表元素的总数为 n 个。控制设备还将第二数据块的块元数据插入至BIL的列表头部,这样能够保证第二数据块被下一次重新访问时,具有更高的概率被替换回LEL中。
- [0151] 参考图7,当第一数据块在BIL链表中命中时,若第二数据块满足第二懒惰条件时,将第一数据块的块元数据插入BIL的列表头部;若第二数据块不满足第二懒惰条件时,将第一数据块的块元数据插入LEL的列表头部,并将第二数据块的块元数据从LEL的列表尾部移动至BIL的列表头部。
- [0152] 步骤420,更新相关数据块的块元数据;
- [0153] 在每一次数据块的访问请求结束后,需要更新相关数据块的块元数据。更新的内容包括但不限于:全部或部分数据块的重用距离、位于缓存空间中的各个数据块的停留时长、第一数据块的flag值、第二数据块的flag值中的至少一种。
- [0154] 可选地,若停留时长是采用上一次访问时间来间接表示时,仅更新第一数据块的上一次访问时间即可。
- [0155] 步骤421,所有访问请求是否已处理完毕;
- [0156] 控制设备检测请求列表中的每个访问请求是否已经处理完毕。
- [0157] 若已经处理完毕,则结束;若请求列表中还存在未处理的访问请求,则进入步骤

422。

[0158] 步骤422,在请求列表中获取下一个访问请求。

[0159] 控制设备从请求列表中获取下一个访问请求,重新进入步骤402进行处理。

[0160] 如图8所示,综上所述,本实施例提供的方法,通过提供两个懒惰条件来进行数据块的重新访问可能性判断,当第一数据块存在于BIL时,第一数据块具有更高的概率被换入缓存空间中,从而使得更有价值的数据块存入缓存空间中,减少缓存空间中的数据块被替换出的可能性,从而提高缓存空间的利用率,也减少对缓存空间的写入次数。

[0161] 本实施例提供的方法,通过在第一数据块为首次访问块且第二数据块符合懒惰条件时,将第一数据块的块元数据移入BIL中,使得第一数据块在下一次被访问时,具有更高的概率被替换入缓存空间中。

[0162] 本实施例提供的方法,通过在第一数据块属于BIL且第二数据块不符合懒惰条件时,将第二数据块的块元数据从缓存空间中删除后移动至BIL的列表头部中进行存储,使得该第二数据块在下一次被访问时,具有更高的概率被替换入缓存空间中。

[0163] 对于本申请的各个实施例,通过实验结果显示,上述缓存方法能极大的提高算法的命中率,从而提升存储系统的读写性能。同时,还能够减少对缓存空间的写入次数,从而延长缓存空间的使用寿命。

[0164] 在一个示意性的例子中,通过采集块级网络存储服务(Cloud Block Service, CBS)真实用户数据,采用日志驱动的试验方法,测试结果显示该算法相对最近最少使用(Least recently used, LRU)算法提升23.80%的命中率,相对读缓存淘汰(Adjustable Replacement Cache, ARC)算法提升15.70%的命中率,相对ARC算法提升9.66%的命中率。

[0165] 在一个示意性的例子中,由于懒惰替换的特性,该方法能够极大的减少对非易失性缓存设备的写入,从而能够有效延长缓存设备的使用寿命。测试结果显示该算法相对LRU算法减少62.85%的缓存设备写入流量,相对ARC算法减少61.85%的缓存设备写入流量。

[0166] 在一个示意性的例子中,参考图8,当用户请求访问数据块A,当数据块A存在于LEL中时,直接从缓存空间84中访问该数据块A。当数据块不存在于LEL中时,检测该数据块是否存在于BIL,以及位于缓存空间84中作为备选剔除块的数据块B是否满足懒惰条件。

[0167] 当数据块A未存在于BIL,且数据块B满足第一懒惰条件时,从存储空间82访问该数据块A。

[0168] 当数据块A存在于BIL,且数据块B满足第二懒惰条件时,从存储空间82访问该数据块A。

[0169] 当数据块A未存在于BIL,且数据块B不满足第一懒惰条件时,从存储空间82读取该数据块A,并使用数据块A在缓存空间84中替换数据块B。

[0170] 当数据块A存在于BIL,且数据块B不满足第一懒惰条件时,从存储空间82读取该数据块A,并使用数据块A在缓存空间84中替换数据块B。

[0171] 以下为本申请的装置实施例,对于装置实施例中未详细描述的细节,可以参考上述方法实施例中的相应记载。

[0172] 图9是本申请一个示例性实施例提供的数据块的缓存装置的结构框图。该缓存装置可以实现成为存储系统的一部分,也可以应用于上述存储系统中,比如实现为控制设备中的一部分。该装置包括:

[0173] 获取模块920,用于获取访问请求,所述访问请求用于请求访问第一数据块;

[0174] 检测模块940,用于当所述第一数据块在所述缓存空间中未命中时,检测第二数据块是否满足懒惰条件,所述第二数据块是所述缓存空间中的备选剔除块,所述懒惰条件是根据重新访问概率确定将所述第二数据块延迟替换出所述缓存空间的条件;

[0175] 访问模块960,用于当所述第二数据块满足所述懒惰条件时,从所述存储空间访问所述第一数据块且不将所述第二数据块替换出所述缓存空间中。所述访问模块960,还用于当所述第二数据块不满足所述懒惰条件时,使用第一数据块替换所述第二数据块。

[0176] 在一个可选的实施例中,所述存储系统存储有BIL,所述BIL存储有历史访问的n个数据块的块元数据,n为正整数。

[0177] 所述检测模块940,用于当所述第一数据块不存在于所述BIL中时,检测所述第二数据块是否满足第一懒惰条件;当所述第一数据块存在于所述BIL中时,检测所述第二数据块是否满足第二懒惰条件;

[0178] 其中,所述第一懒惰条件所指示的重新访问概率小于所述第二懒惰条件所指示的重新访问概率。

[0179] 在一个可选的实施例中,所述第一懒惰条件包括:所述第二数据块的所述标记值大于第一阈值,所述标记值是用于衡量所述第二数据块的访问重要程度的变量;所述第二懒惰条件包括:所述第二数据块的所述标记值大于第二阈值,且所述第二数据块在所述缓存空间中的停留时长大于K倍的块平均重用距离。

[0180] 在一个可选的实施例中,当所述第二数据块在所述缓存空间中命中时,按照第一幅度增加所述标记值;当所述第二数据块在所述缓存空间中因满足所述懒惰条件未被替换时,按照第二幅度减少所述标记值。

[0181] 在一个可选的实施例中,所述装置还包括:元数据修改模块980,如图9所示。

[0182] 所述元数据修改模块980,用于当所述第二数据块满足所述第一懒惰条件时,将所述第一数据块的块元数据添加至所述BIL的列表头部,并将所述BIL的列表尾部的第三数据块的块元数据进行删除,以及当所述第二数据块满足所述第二懒惰条件时,将所述第一数据块的块元数据移动至所述BIL的列表头部。

[0183] 在一个可选的实施例中,所述存储系统存储有LEL,所述LEL存储有位于所述缓存空间中的n个数据块的块元数据,n为正整数。

[0184] 当所述第二数据块不满足所述第一懒惰条件时,所述元数据修改模块980用于将所述第二数据块的块元数据从所述LEL的列表尾部删除。

[0185] 当所述第二数据块不满足所述第二懒惰条件时,所述元数据修改模块980用于将所述第二数据块的块元数据从所述LEL的列表尾部删除,且将所述第二数据块的块元数据移动至所述BIL的列表头部。

[0186] 当所述第二数据块不满足所述懒惰条件时,所述元数据修改模块980用于将所述第一数据块的块元数据移动至所述LEL的列表头部。

[0187] 可选地,所述懒惰条件包括第一懒惰条件和第二懒惰条件。

[0188] 图11是本申请一个实施例提供的计算机设备的结构示意图。具体来讲:计算机设备1000包括中央处理单元(英文:Central Processing Unit,简称:CPU)1001、包括随机存取存储器(英文:random access memory,简称:RAM)1002和只读存储器(英文:read-only

memory, 简称:ROM) 1003 的系统存储器 1004, 以及连接系统存储器 1004 和中央处理单元 1001 的系统总线 1005。所述计算机设备 1000 还包括帮助计算机内的各个器件之间传输信息的基本输入/输出系统 (I/O 系统) 1006, 和用于存储操作系统 1013、应用程序 1014 和其他程序模块 1015 的大容量存储设备 1007。

[0189] 所述基本输入/输出系统 1006 包括有用于显示信息的显示器 1008 和用于用户输入信息的诸如鼠标、键盘之类的输入设备 1009。其中所述显示器 1008 和输入设备 1009 都通过连接到系统总线 1005 的输入/输出控制器 1010 连接到中央处理单元 1001。所述基本输入/输出系统 1006 还可以包括输入/输出控制器 1010 以用于接收和处理来自键盘、鼠标、或电子触控笔等多个其他设备的输入。类似地, 输入/输出控制器 1010 还提供输出到显示屏、打印机或其他类型的输出设备。

[0190] 所述大容量存储设备 1007 通过连接到系统总线 1005 的大容量存储控制器 (未示出) 连接到中央处理单元 1001。所述大容量存储设备 1007 及其相关联的计算机可读介质为计算机设备 1000 提供非易失性存储。也就是说, 所述大容量存储设备 1007 可以包括诸如硬盘或者只读光盘 (英文: Compact Disc Read-Only Memory, 简称: CD-ROM) 驱动器之类的计算机可读介质 (未示出)。

[0191] 不失一般性, 所述计算机可读介质可以包括计算机存储介质和通信介质。计算机存储介质包括以用于存储诸如计算机可读指令、数据结构、程序模块或其他数据等信息的任何方法或技术实现的易失性和非易失性、可移动和不可移动介质。计算机存储介质包括 RAM、ROM、可擦除可编程只读存储器 (英文: erasable programmable read-only memory, 简称: EPROM)、电可擦除可编程只读存储器 (英文: electrically erasable programmable read-only memory, 简称: EEPROM)、闪存或其他固态存储其技术, CD-ROM、数字通用光盘 (英文: Digital Versatile Disc, 简称: DVD) 或其他光学存储、磁带盒、磁带、磁盘存储或其他磁性存储设备。当然, 本领域技术人员可知所述计算机存储介质不局限于上述几种。上述的系统存储器 1004 和大容量存储设备 1007 可以统称为存储器。

[0192] 根据本申请的各种实施例, 所述计算机设备 1000 还可以通过诸如因特网等网络连接到网络上的远程计算机运行。也即计算机设备 1000 可以通过连接在所述系统总线 1005 上的网络接口单元 1011 连接到网络 1012, 或者说, 也可以使用网络接口单元 1011 来连接到其他类型的网络或远程计算机系统 (未示出)。

[0193] 其中, 上述处理器被配置为加载并执行可执行指令以实现如上述方法实施例提供的数据块的缓存方法。

[0194] 本申请还提供一种计算机可读存储介质, 所述存储介质中存储有至少一条指令、至少一段程序、代码集或指令集, 所述至少一条指令、所述至少一段程序、所述代码集或指令集由所述处理器加载并执行以实现上述方法实施例提供的数据块的缓存方法。

[0195] 本申请还提供了一种计算机程序产品, 当其在电子设备上运行时, 使得电子设备执行上述各个方法实施例所述的数据块的缓存方法。

[0196] 上述本申请实施例序号仅仅为了描述, 不代表实施例的优劣。

[0197] 本领域普通技术人员可以理解实现上述实施例的全部或部分步骤可以通过硬件来完成, 也可以通过程序来指令相关的硬件完成, 所述的程序可以存储于一种计算机可读存储介质中, 上述提到的存储介质可以是只读存储器, 磁盘或光盘等。

[0198] 以上所述仅为本申请的较佳实施例,并不用以限制本申请,凡在本申请的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本申请的保护范围之内。

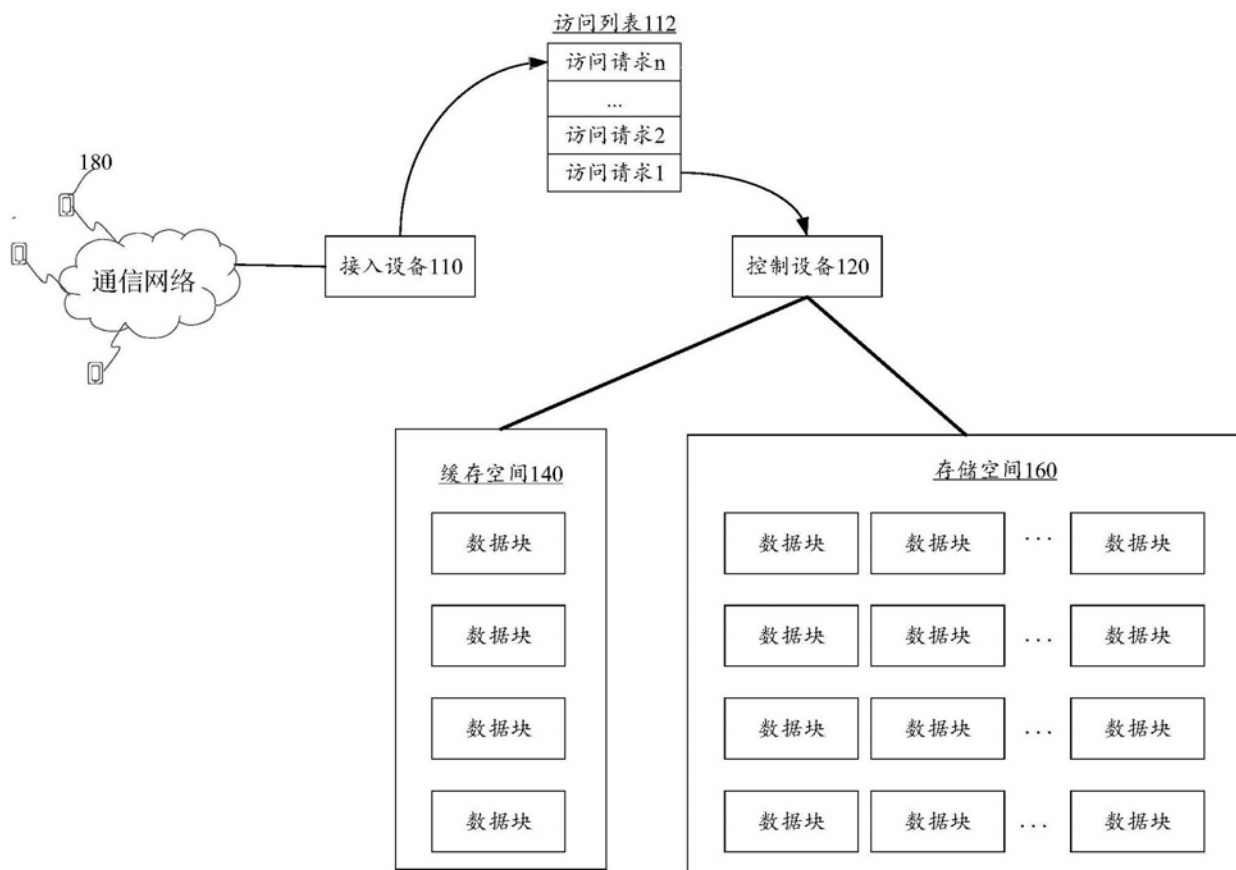


图1

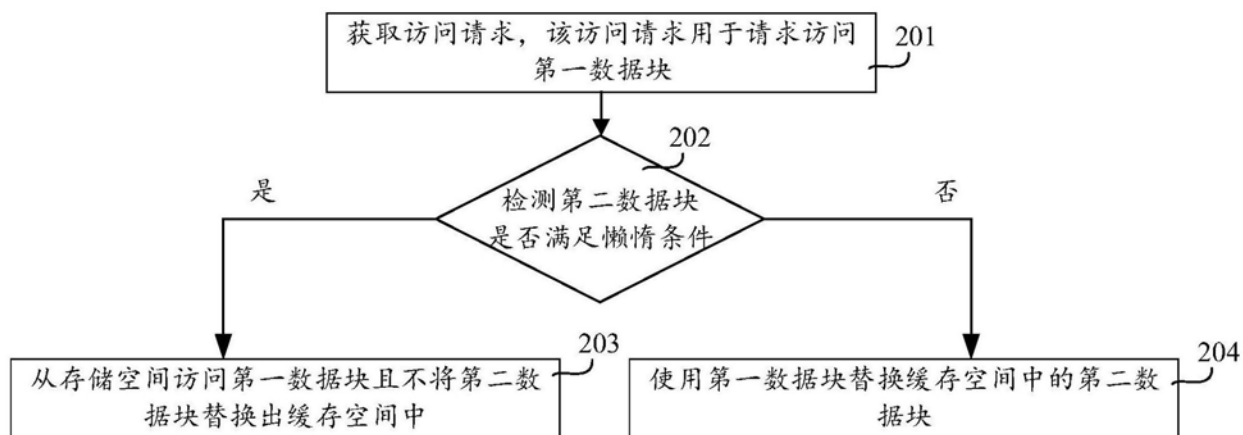


图2

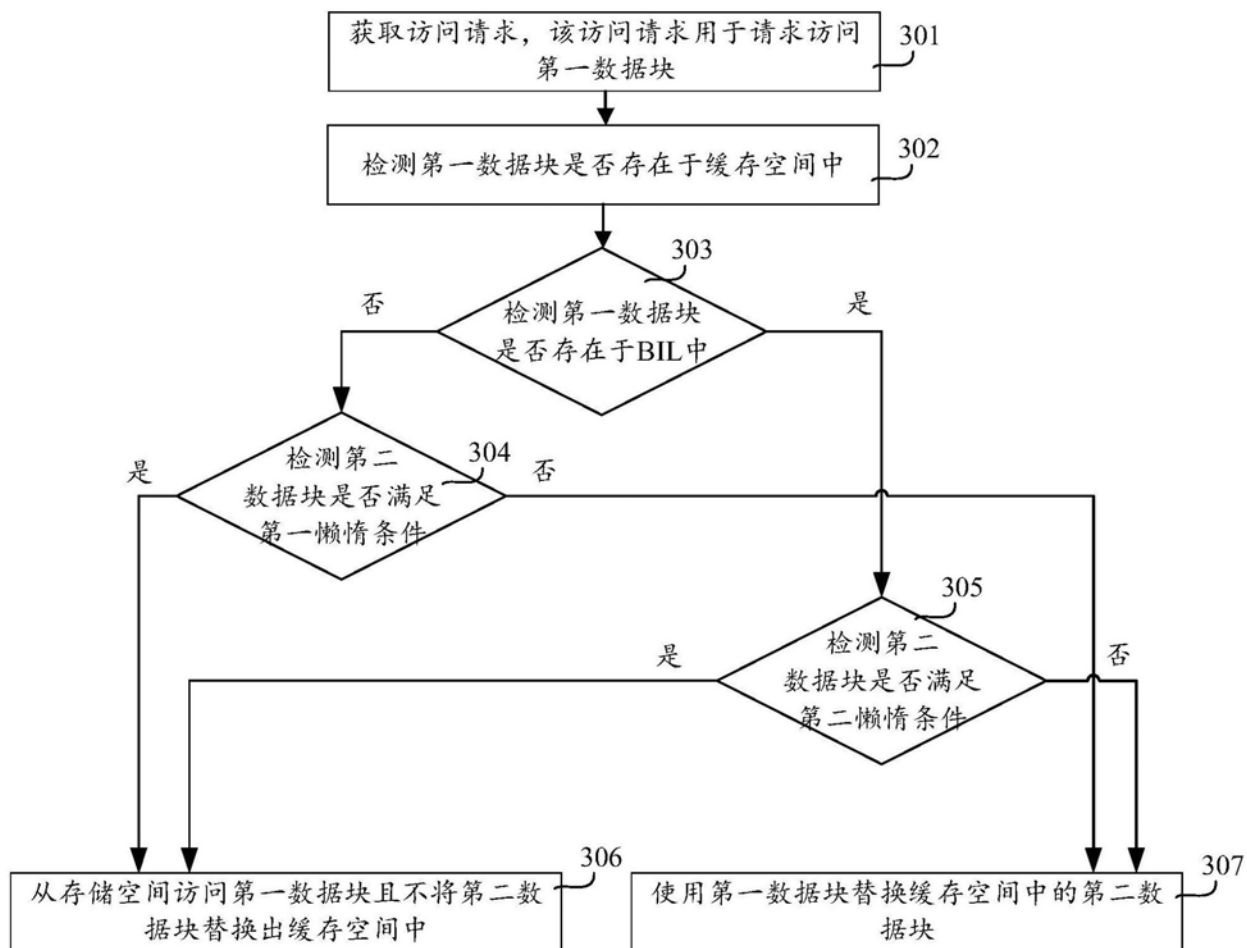


图3

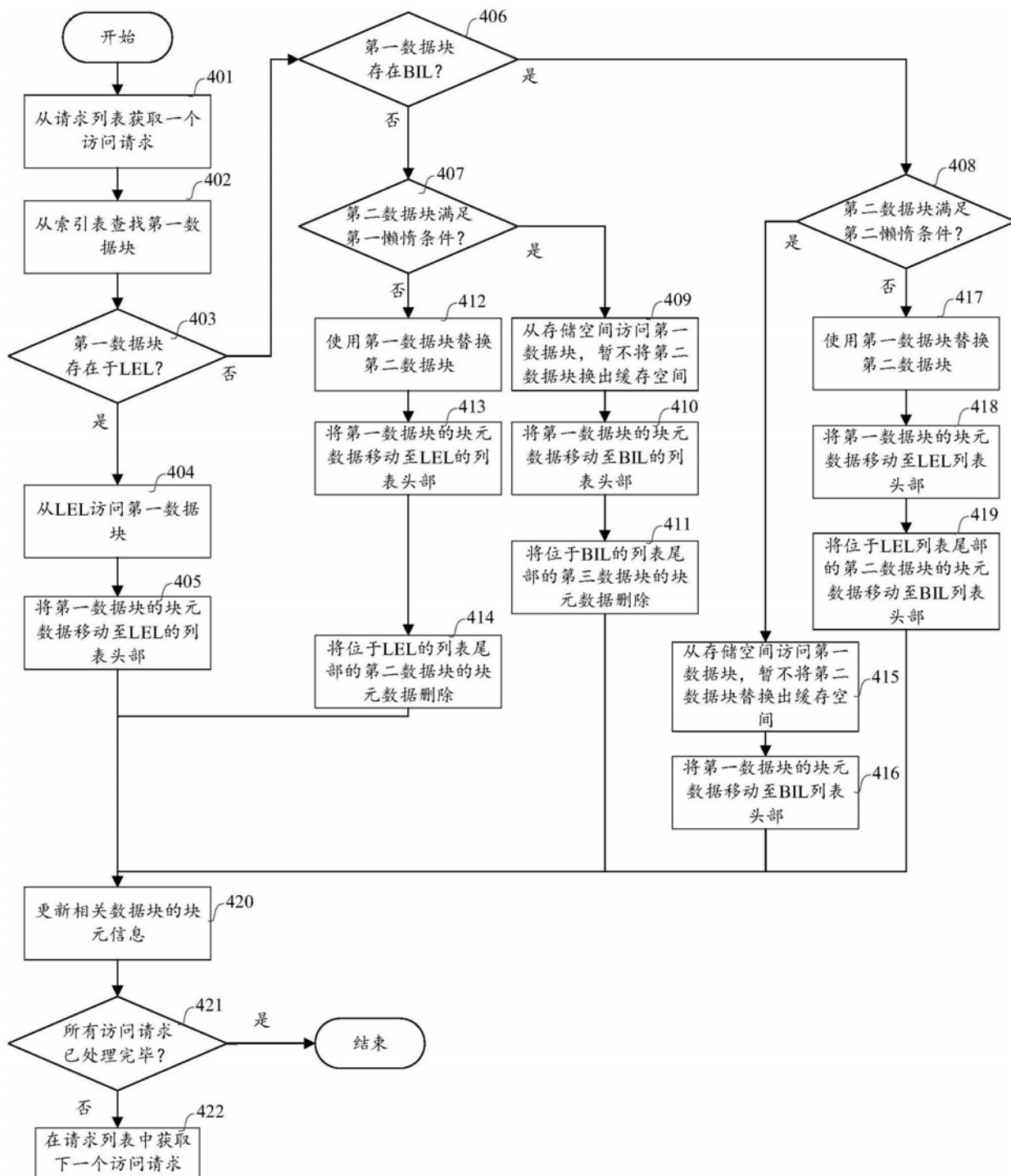


图4

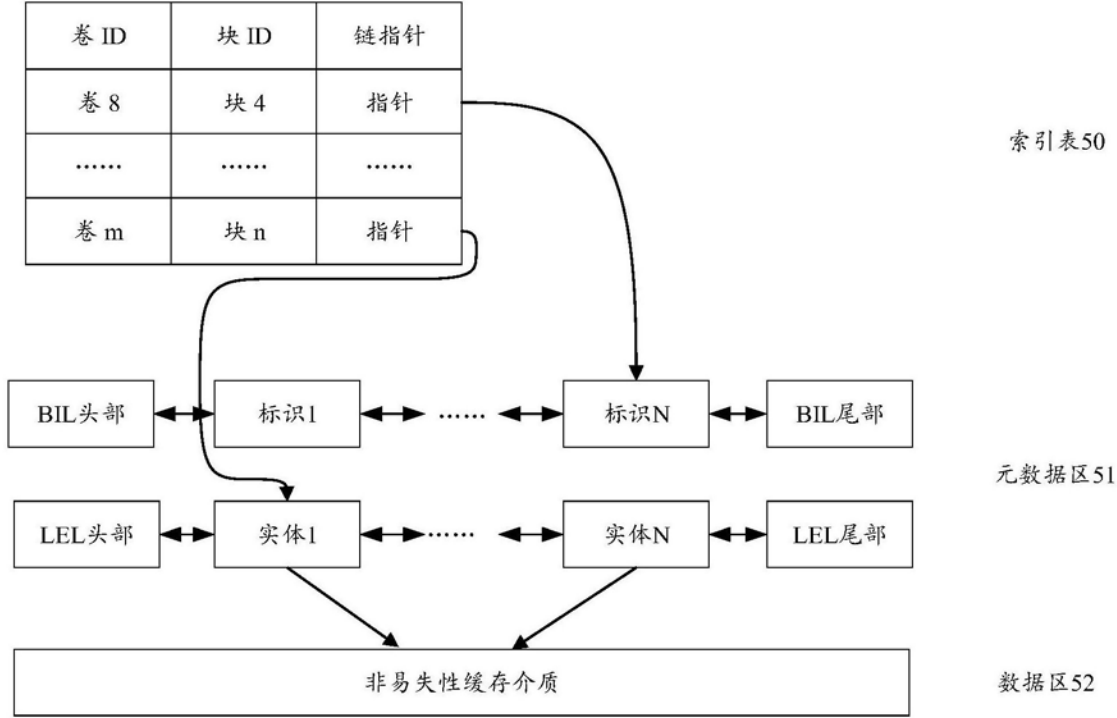


图5

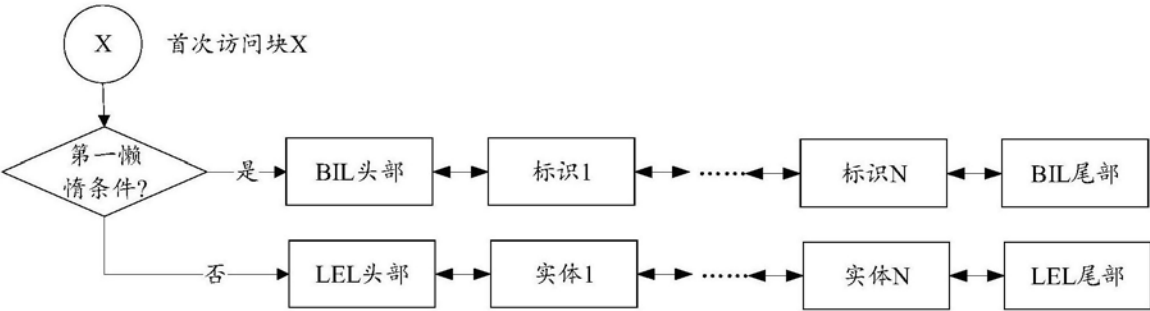


图6

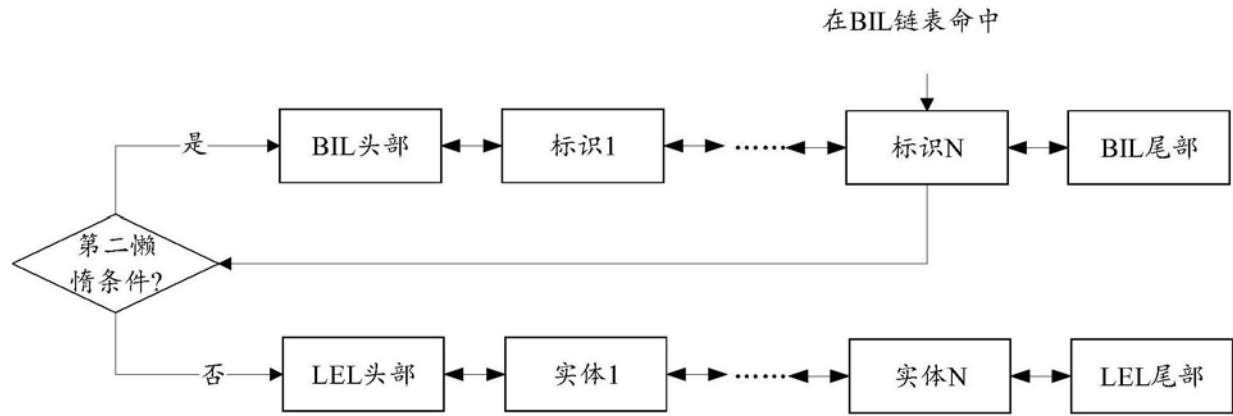


图7

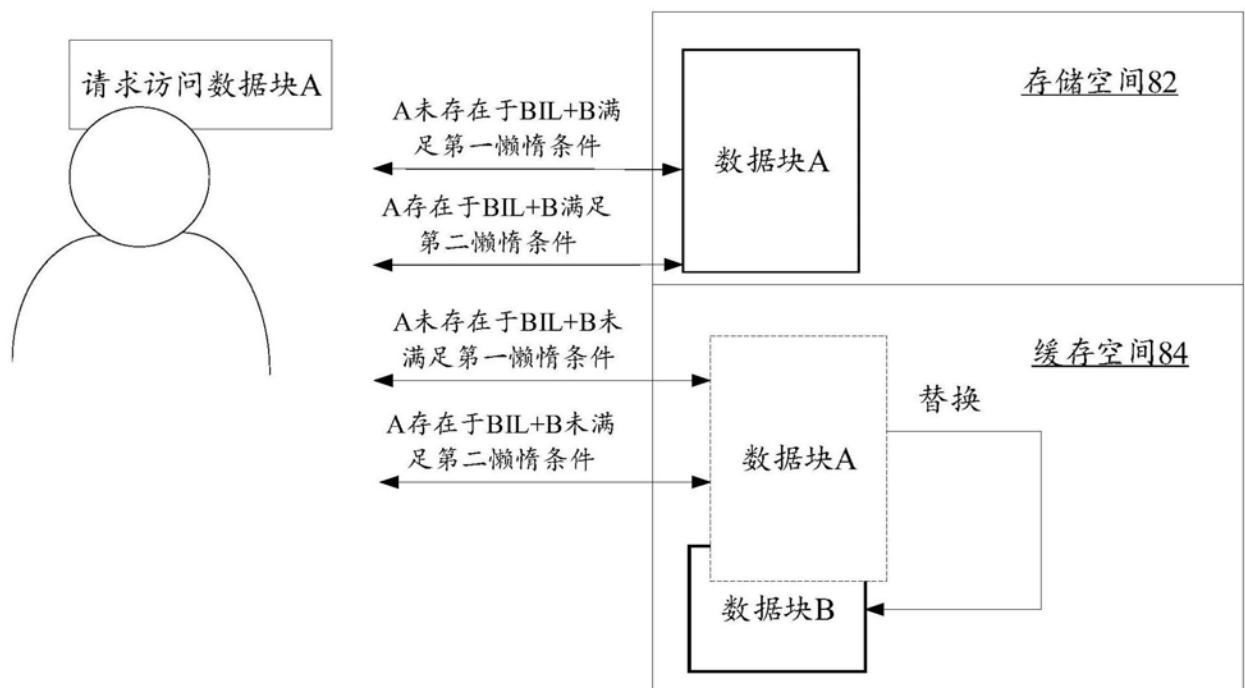


图8

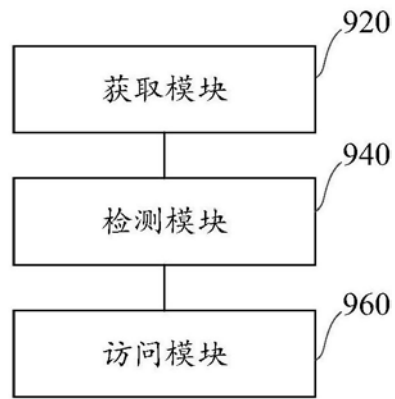


图9

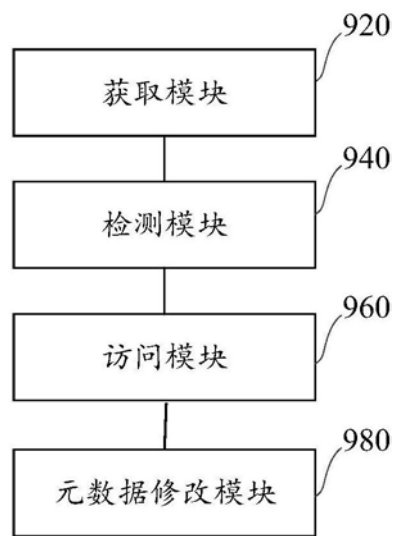


图10

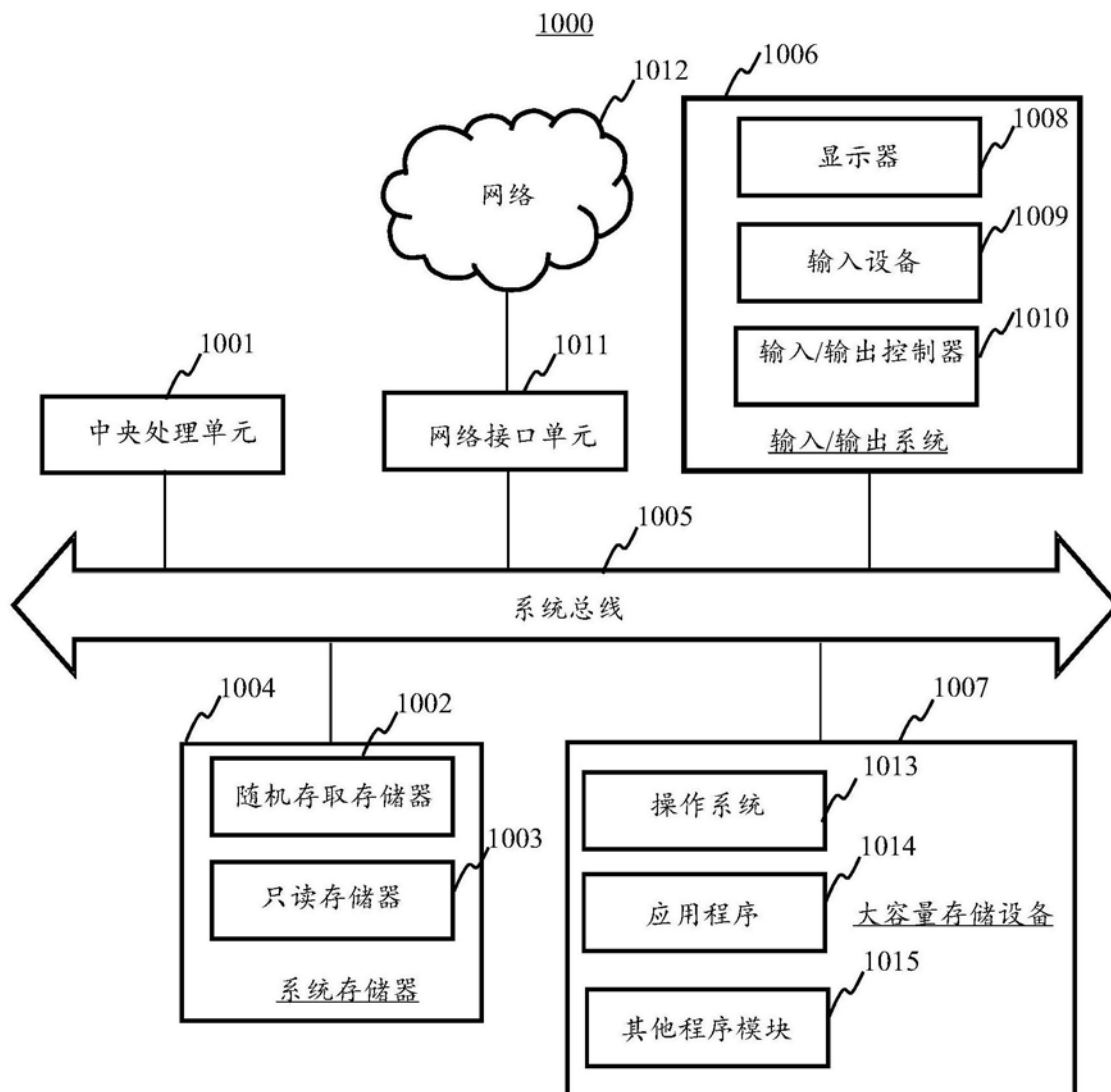


图11