



(12) 发明专利申请

(10) 申请公布号 CN 105009093 A

(43) 申请公布日 2015. 10. 28

(21) 申请号 201480009440. 1

(51) Int. Cl.

(22) 申请日 2014. 02. 12

G06F 12/02(2006. 01)

G06F 12/08(2006. 01)

(30) 优先权数据

13/772, 400 2013. 02. 21 US

(85) PCT国际申请进入国家阶段日

2015. 08. 19

(86) PCT国际申请的申请数据

PCT/US2014/015994 2014. 02. 12

(87) PCT国际申请的公布数据

W02014/130317 EN 2014. 08. 28

(71) 申请人 高通股份有限公司

地址 美国加利福尼亚州

(72) 发明人 X·董

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 袁逸

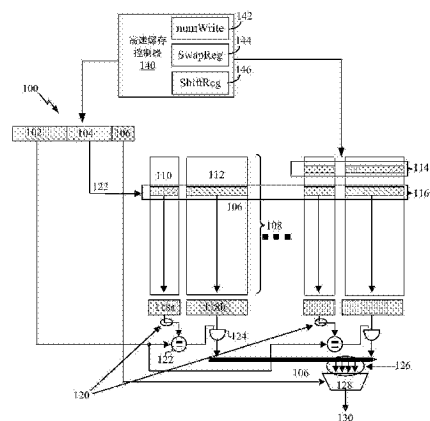
权利要求书3页 说明书10页 附图8页

(54) 发明名称

对具有有限耐写性的高速缓存的集间损耗矫平

(57) 摘要

高速缓存控制器包括第一寄存器,该第一寄存器在对高速缓存存储器中的数个高速缓存集的每一次存储器位置交换操作之后更新,并且每N-1次存储器位置交换操作就复位。N是该高速缓存存储器中的高速缓存集的数目。该存储器控制器还具有第二寄存器,该第二寄存器在每N-1次存储器位置交换操作后更新,并且每(N<sup>2</sup>-N)次存储器位置交换操作就复位。第一和第二寄存器跟踪这些高速缓存集的逻辑位置和物理位置之间的关系。



1. 一种用于对高速缓存存储器进行损耗矫平的方法,包括:

通过对所述高速缓存存储器的存储器写操作的数目达到阈值时对高速缓存集执行多个存储器位置交换操作来动态地轮转所述高速缓存存储器的高速缓存集,每个交换操作包括仅清除来自经交换的高速缓存集的内容,而保持其他高速缓存集的存储器内容完好;以及

跟踪所述经交换的高速缓存集以将逻辑高速缓存集号转换成物理高速缓存集号。

2. 如权利要求 1 所述的方法,其特征在于,动态地轮转所述高速缓存集进一步包括:在全局计数器中对向所述高速缓存存储器所作的存储器写操作的数目进行计数;当所述全局计数器达到所述阈值时发起所述多个存储器位置交换操作;以及将所述全局计数器复位。

3. 如权利要求 2 所述的方法,其特征在于,动态地轮转所述高速缓存集包括:

将交换寄存器 (SwapReg) 设置成  $(\text{SwapReg}+1)\text{mod}(N-1)$ , 其中  $N$  是所述高速缓存存储器中的所述高速缓存集的数目;以及

当所述 SwapReg 为零时将移位寄存器 (ShiftReg) 设置成  $(\text{ShiftReg}+1)\text{mod}N$ 。

4. 如权利要求 1 所述的方法,其特征在于,进一步包括:

用全局计数器通过对于向所述高速缓存存储器所作的每次写操作递增所述全局计数器来对向所述高速缓存存储器所作的存储器写操作的数目进行计数;

当所述全局计数器等于所述阈值时执行交换操作并递增交换计数器,其中执行所述交换操作包括重映射经交换的高速缓存集;

当交换计数器值等于  $N-1$  时,递增移位计数器并且将所述交换计数器值复位,其中  $N$  是所述高速缓存存储器中的所述高速缓存集的数目;以及

当移位计数器值等于  $N$  时,将所述移位计数器复位。

5. 如权利要求 4 所述的方法,其特征在于,重映射所述经交换的高速缓存集包括:

将所述经交换的高速缓存集内的脏数据存储到写回缓冲器中;以及

将所述经交换的高速缓存集中的其他数据无效化。

6. 如权利要求 4 所述的方法,其特征在于,将输入高速缓存集号转换成输出高速缓存集号包括:

当输入逻辑集号等于所述交换计数器值时,将输出物理集号设置成所述移位计数器值;

当所述输入逻辑集号大于所述交换计数器值时,将所述输出物理集号设置成 (所述输入逻辑集号 + 所述移位计数器值) mod  $N$ ; 以及

当所述输入逻辑集号小于所述交换计数器值时,将所述输出物理集号设置成 (所述输入逻辑集号 + 所述移位计数器值 + 1) mod  $N$ 。

7. 如权利要求 4 所述的方法,其特征在于,跟踪所述经交换的高速缓存集包括:

当输入物理集号等于所述移位计数器值时,将输出逻辑集号设置成所述交换计数器值;

当 (所述输入物理集号 - 所述移位计数器值) mod  $N$  大于所述交换计数器值时,将所述输出逻辑集号设置成 (所述输入物理集号 - 所述移位计数器值) mod  $N$ ; 以及

否则,将所述输出逻辑集号设置成 (所述输入物理集号 - 所述移位计数器值 - 1) mod  $N$ 。

8. 如权利要求 1 所述的方法,其特征在于,进一步包括:将所述高速缓存存储器集成到移动电话、机顶盒、音乐播放器、视频播放器、娱乐单元、导航设备、计算机、手持式个人通信系统 (PCS) 单元、便携式数据单元、和 / 或固定位置数据单元中。

9. 一种高速缓存控制器,包括:

第一寄存器,其在对高速缓存存储器的多个高速缓存集的每次存储器位置交换操作后更新,并且每  $N-1$  次存储器交换操作后复位,其中  $N$  是所述高速缓存存储器中的高速缓存集的数目;以及

第二寄存器,其在对所述高速缓存存储器的所述多个高速缓存集的每  $N-1$  次存储器位置交换操作后更新,并且每  $(N^2-N)$  次存储器位置交换操作就复位,其中所述第一寄存器和所述第二寄存器跟踪所述高速缓存集的逻辑位置与物理位置之间的关系。

10. 如权利要求 9 所述的高速缓存控制器,其特征在于,所述高速缓存控制器能操作以动态地交换两个存储器位置并且跟踪所述高速缓存集的所述交换操作以使用所述第一寄存器和所述第二寄存器将逻辑高速缓存集号转换成物理高速缓存集号。

11. 如权利要求 9 所述的高速缓存控制器,其特征在于,所述第一寄存器是交换寄存器 (SwapReg),其在执行存储器交换操作时被设置成  $(\text{SwapReg}+1) \bmod (N-1)$ 。

12. 如权利要求 11 所述的高速缓存控制器,其特征在于,所述第二寄存器是移位寄存器 (ShiftReg),其在所述 SwapReg 等于零时被设置成  $(\text{ShiftReg}+1) \bmod N$ 。

13. 如权利要求 9 所述的高速缓存控制器,其特征在于

所述第一寄存器是交换寄存器 (SwapReg),其在执行存储器交换操作时被设置成  $(\text{SwapReg}+1) \bmod (N-1)$ ,并且

其中所述第二寄存器是移位寄存器 (ShiftReg),其在所述 SwapReg 等于零时被设置成  $(\text{ShiftReg}+1) \bmod N$  的。

14. 如权利要求 13 所述的高速缓存控制器,其特征在于,所述高速缓存控制器能操作以动态地交换两个存储器位置并且跟踪所述高速缓存集的所述交换操作以使用所述 SwapReg 和所述 ShiftReg 来将逻辑高速缓存集号转换成物理高速缓存集号。

15. 如权利要求 9 所述的高速缓存控制器,其特征在于,所述高速缓存控制器能操作以在写计数器寄存器等于阈值时发起所述存储器位置交换操作。

16. 如权利要求 9 所述的高速缓存控制器,其特征在于,所述高速缓存控制器被集成到移动电话、机顶盒、音乐播放器、视频播放器、娱乐单元、导航设备、计算机、手持式个人通信系统 (PCS) 单元、便携式数据单元、和 / 或固定位置数据单元中。

17. 一种用于对高速缓存存储器进行损耗矫平的高速缓存控制器,包括:

用于通过在对所述高速缓存存储器的存储器写操作的数目达到阈值时对高速缓存集执行多个存储器位置交换操作来动态地轮转所述高速缓存存储器的高速缓存集的装置,每个交换操作包括仅清除来自经交换的高速缓存集的内容而保持其他高速缓存集的存储器内容完好;以及

用于跟踪所述经交换的高速缓存集以将逻辑高速缓存集号转换成物理高速缓存集号的装置。

18. 如权利要求 17 所述的高速缓存控制器,其特征在于,所述高速缓存控制器被集成到移动电话、机顶盒、音乐播放器、视频播放器、娱乐单元、导航设备、计算机、手持式个人通

信系统 (PCS) 单元、便携式数据单元、和 / 或固定位置数据单元中。

19. 一种用于对高速缓存存储器进行损耗矫平的方法, 包括:

通过对所述高速缓存存储器的存储器写操作的数目达到阈值时对高速缓存集执行多个存储器位置交换操作来动态地轮转所述高速缓存存储器的高速缓存集的步骤, 每个交换操作包括仅清除来自经交换的高速缓存集的内容而保持其他高速缓存集的存储器内容完好; 以及

跟踪所述经交换的高速缓存集以将逻辑高速缓存集号转换成物理高速缓存集号的步骤。

20. 如权利要求 19 所述的方法, 其特征在于, 进一步包括: 将所述高速缓存存储器集成到移动电话、机顶盒、音乐播放器、视频播放器、娱乐单元、导航设备、计算机、手持式个人通信系统 (PCS) 单元、便携式数据单元、和 / 或固定位置数据单元中。

## 对具有有限耐写性的高速缓存的集间损耗矫平

### 技术领域

[0001] 本公开一般涉及存储器和高速缓存。更具体地,本公开涉及对具有有限耐写性的高速缓存的集间损耗矫平。

### 背景技术

[0002] 对于可被用于无线通信或其他应用的高速数字电子器件,使用了非易失性存储器。然而,非易失性存储器(诸如电阻性随机存取存储器(ReRAM)和相变随机存取存储器(PCRAM))具有有限的耐写性。耐写性可以被定义为在存储介质变得不可靠之前可被施加于存储器块的程序/循环的数目,并且一般是通过估计该存储器被用得有多频繁以及多彻底来计算的。换句话说,耐写性衡量某个类型的存储介质的服务寿命。

[0003] 损耗矫平是用以延长存储介质的耐写性(例如,服务寿命)的技术,并且是高速缓存设计的一部分。一种损耗矫平办法安排数据以使得重写被均匀地跨整个存储介质分布。以这种方式,不会有单个块由于写循环的高度集中而发生故障。其他损耗矫平的办法可以包括每次写发生时就动态更新映射,该映射随后将被写块链接到新块。另一个办法静态地保持这些块相同而不替换它们,但是周期性地轮转这些块从而使得这些块可以被其他数据使用。

[0004] 对非易失性存储器(例如,其也可被用在计算机的主存储器中)的损耗矫平是公知的并且已经得到很好的探索。但是,当对片上高速缓存使用损耗矫平时,通常对非易失性存储器采用的传统损耗矫平办法显现出太多的性能开销。因此,高性能开销抑制了损耗矫平技术对具有有限耐写性的高速缓存的有效性。

### [0005] 概述

[0006] 根据本公开的一方面,描述了用以对高速缓存存储器进行集间损耗矫平的高速缓存控制器。该高速缓存控制器包括第一寄存器,该第一寄存器在对该高速缓存存储器的高速缓存集的每一次存储器位置交换操作之后更新,并且每  $N-1$  次存储器位置交换操作就复位。 $N$  是该高速缓存存储器中的高速缓存集的数目。该高速缓存控制器进一步包括第二寄存器,该第二寄存器在对该高速缓存存储器的高速缓存集的每  $N-1$  次存储器位置交换操作之后更新,并且每  $(N^2-N)$  次存储器位置交换操作就复位。第一和第二寄存器可跟踪这些高速缓存集的逻辑位置和物理位置之间的关系。

[0007] 根据本公开的另一个方面,描述了一种用于对高速缓存存储器进行集间损耗矫平的方法。该方法包括通过在对该高速缓存存储器的存储器写操作的次数达到阈值时对高速缓存集执行存储器位置交换操作来动态地轮转该高速缓存存储器的这些高速缓存集。每个交换操作可以包括仅清除来自经交换高速缓存集的内容,而保持其他高速缓存集的存储器内容完好。该方法还包括跟踪这些经交换高速缓存集以将逻辑高速缓存集号转换成物理高速缓存集号。

[0008] 根据本公开的进一步方面,描述了用以对高速缓存存储器进行集间损耗矫平的高速缓存控制器。该高速缓存控制器包括用于通过在对高速缓存存储器的存储器写操作的次

数达到阈值时对高速缓存集执行存储器位置交换操作来动态地轮转该高速缓存存储器的  
高速缓存集的装置。每个交换操作可以包括仅清除来自经交换高速缓存集的内容,而保持  
其他高速缓存集的存储器内容完好。该高速缓存控制器进一步包括用于跟踪这些经交换高  
速缓存集以将逻辑高速缓存集号转换成物理高速缓存集号的装置。

[0009] 根据本公开的另一个方面,描述了一种用于对高速缓存存储器进行集间损耗矫平  
的方法。该方法包括通过在该高速缓存存储器的存储器写操作的次数达到阈值时对高速  
缓存集执行存储器位置交换操作来动态地轮转该高速缓存存储器的高速缓存集的步骤。每  
个交换操作可以包括仅清除来自经交换高速缓存集的内容,而保持其他高速缓存集的存储  
器内容完好。该方法还包括跟踪这些经交换高速缓存集以将逻辑高速缓存集号转换成物理  
高速缓存集号的步骤。

[0010] 这已较宽泛地勾勒出本公开的特征和技术优势以便下面的详细描述可以被更好  
地理解。本公开的其他特征和优点将在下文描述。本领域技术人员应该领会,本公开可容  
易地用作修改或设计用于实施与本公开相同的目的的其他结构的基础。本领域技术人员  
还应认识到,这样的等效构造并不脱离所附权利要求中所阐述的本公开的教导。被认为是  
本公开的特性的新颖特征在其组织和操作方法两方面连同进一步的目的和优点在结合附  
图来考虑以下描述时将被更好地理解。然而,要清楚理解的是,提供每一幅附图均仅用于解  
说和描述目的,且无意作为对本公开的限定的定义。

[0011] 附图简述

[0012] 为了更全面地理解本公开,现在结合附图参阅以下描述。

[0013] 图 1 是根据本公开的一方面的示例高速缓存存储器的示图,该示例高速缓存存储  
器包括用于该高速缓存存储器的损耗矫平的高速缓存控制器。

[0014] 图 2 是示出根据本公开的一方面的用于损耗矫平的示例轮转的示图。

[0015] 图 3 是示出根据本公开的一方面的损耗矫平的操作的逻辑流程图。

[0016] 图 4 是示出根据本公开的一方面的损耗矫平的操作的逻辑流程图。

[0017] 图 5 是解说根据本公开的一方面的用于损耗矫平的方法的过程流程图。

[0018] 图 6 是解说根据本公开的一方面的用于损耗矫平的方法的过程流程图。

[0019] 图 7 是示出其中可有利地采用本公开的配置的示例性无线通信系统的框图。

[0020] 图 8 是解说根据一种配置的用于半导体组件的电路、布局、以及逻辑设计的设计  
工作站的框图。

[0021] 详细描述

[0022] 以下结合附图阐述的详细描述旨在作为各种配置的描述,而无意表示可实践本文  
中所描述的概念的仅有的配置。本详细描述包括具体细节以便提供对各种概念的透彻理  
解。然而,对于本领域技术人员将显而易见的是,没有这些具体细节也可实践这些概念。在  
一些实例中,以框图形式示出众所周知的结构和组件以避免湮没此类概念。如本文所述  
的,术语“和/或”的使用旨在代表“可兼性或”,而术语“或”的使用旨在代表“排他性或”。

[0023] 存储器(诸如静态随机存取(SRAM)和嵌入式动态RAM(eDRAM))通常被用于现代  
微处理器中的片上高速缓存设计。现代的计算机和设备还规定更大的片上高速缓存,但是  
传统SRAM或eDRAM高速缓存的可伸缩性正日益受技术局限性(诸如,泄漏功率和单元密  
度)所束缚。近来,新的非易失性存储器(NVM)技术(诸如举例而言,相变随机存取存储器

(RAM)、自旋转移矩 RAM、以及电阻性 RAM) 已经被探索作为有望用于片上高速缓存的替代性存储器技术。相比于传统存储器 (诸如 SRAM 和 eDRAM), 这些新兴的非易失性存储器技术具有高密度、低待机功率、低电压、更好可伸缩性以及非易失性的共同优势。然而, 对它们的采用受到它们有限的耐写性的妨碍。这个问题将会被现有的高速缓存管理策略所放大, 结果导致高速缓存块上失衡的写业务, 因为此类策略并不知晓写差异。这些策略原先是被设计用于 SRAM 高速缓存的, 并且结果导致在向高速缓存块的写入方面的显著的非均一性, 这使得被繁重地写入的高速缓存块比绝大多数其他块产生故障要快得多或早得多。

[0024] 许多损耗矫平技术已被提议以用于延长非易失性存储器技术的寿命, 但是高速缓存与主存储器操作机制之间的差异使得用于非易失性存储器的现有损耗矫平技术不胜任用于非易失性高速缓存。为了解决这些问题并且减少集间写差异, 提供了交换移位方案以减少非易失性存储器高速缓存的集间写差异。该方案具有非常小的硬件开销, 其仅使用一个全局计数器和两个全局寄存器。通过采用这种方案, 低级片上非易失性存储器高速缓存的寿命能被改善。

[0025] 在使用具有有限耐写性的非易失性存储器来设计任何高速缓存或存储器子系统时, 写差异是一个重要的考虑因素。大的写差异可以很大程度上使得产品寿命降级, 因为只要一小子集的经历了最差情形写业务的存储器单元就能够导致死缓存或存储器子系统, 即便在大部分单元还远没有被用坏时亦是如此。

[0026] 图 1 是根据本公开的一方面的示例性高速缓存存储器 100 的示图, 该高速缓存存储器 100 包括用于高速缓存存储器 100 的集间损耗矫平的高速缓存控制器 140。高速缓存存储器 100 包括页号 102、集号 140、字节号 106、高速缓存路 108、标记部分 110、数据部分 112、高速缓存块 114、高速缓存集 116、标记感测放大器 118a、数据感测放大器 118b、标记输出 120、比较器 122、逻辑门 124、高速缓存群 126、选择电路系统 128 和字输出 130。

[0027] 高速缓存存储器 100 中的地址可包括页号 102、集号 104 和字节号 106。在一个实现中, 页号 102 可以是虚拟页号。集号 104 对应于诸高速缓存集 116 之一。高速缓存块 114 包括标记部分 110 和数据部分 112。标记部分 110 可以包含数据部分 112 中的实际数据的地址的一部分、或者其他标识信息以在数据部分 112 中定位数据。数据部分 112 包含实际数据。如可由图 1 中的横向编组看出的, 高速缓存集 116 中的一个是一组高速缓存块 114。如图 1 中所能看出的, 高速缓存路 108 是另一组高速缓存块 114, 但是是纵向编组的。标记感测放大器 118a 和数据感测放大器 118b 感测来自高速缓存条目的逻辑电平, 从而数据在输出时被正确解读 (为 1 或 0)。

[0028] 标记输出 120 处的数据 (其为标记感测放大器 118a 的输出) 可以包含页帧号、有效位和相干性位。可以接着由比较器 122 将来自标记输出 120 的数据与页号 102 作比较, 以查看这两个值是否相等。若这些值相等并且有命中, 那么比较器 122 的输出连同数据感测放大器 118b 的输出一起被输入到逻辑门 124。逻辑门 124 的输出出现在高速缓存群 126 中。在一个实现中, 高速缓存群 126 之一包含多个字。高速缓存群 126 被输入到将字节号 106 用作选择输入的选择电路系统 128。将字节号 106 用作选择输入的选择电路系统 128 的输出是字输出 130。

[0029] 图 1 也是 n 路集关联式高速缓存的示例框图, 并且可能有其他类型的高速缓存用于本公开。集关联式高速缓存可以由并行操作的多个直接映射式高速缓存构成 (例如, 一

个直接映射式高速缓存可以是包括标记部分 110 和数据部分 112 的高速缓存条目)。数据读出可以由与页号、以及块有效位(其可为标记或元数据条目的一部分)和页许可(页号 102 的一部分)的标记比较来控制。高速缓存列大小也可以等于虚拟存储器页大小,并且高速缓存索引可以不使用来自页号 102 或虚拟页号的位。

[0030] 个体的高速缓存块 114 在一个方向上被编组成高速缓存集 116 之一,并且在另一方向上被编组成高速缓存路 108。高速缓存块 114 取决于它们有多频繁地被写入而也可以具有负载分布。一些高速缓存块 114 被繁重地写入,而其他高速缓存块 114 很少被写入。因此,这引发了集间写差异。集间写差异发生于在写活动方面有差异的高速缓存集 116 间。即,来自一个高速缓存集 116 的高速缓存块 114 可能比来自另一高速缓存集 116 的高速缓存块 114 更多或更少地被写入。进一步,当不同高速缓存集 116 相比于其他高速缓存集 116 而言作为整体被写入时,集间写差异也会发生。

[0031] 在该配置中,高速缓存控制器 140 平均地将写业务分布到高速缓存路 108 的不同行和不同列。虽然在图 1 中高速缓存路 108 被示为占据列而高速缓存集 116 被示为占据行,但是高速缓存存储器 100 的实现并不限于该配置。特别地,高速缓存路 108 可占据行或高速缓存存储器 100 内可不为列的其他结构,并且高速缓存集 116 可以占据列或者高速缓存存储器 100 内可不为行的其他结构。

[0032] 主存储器损耗矫平技术通常使用数据移动来实现地址重映射。这是因为在主存储器中,在每次重映射后,数据不可能丢失并且能够被移动到新的位置。然而,数据移动操作总是会招致面积和性能开销。首先,数据移动需要临时数据存储位置以接收该数据。其次,一次高速缓存集移动涉及数次块读和写操作。因此,在数据移动期间,高速缓存端口被阻塞并且系统性能必然降级。当一个示例数据移动方案从主存储器被扩展到应用于存储器高速缓存时,一个附加的高速缓存集(间隔集)被添加并且从一个集合去往该间隔集的数据被周期性移动。因为高速缓存技术是更为性能敏感的,所以不能直接使用主存储器损耗矫平技术。因此,当设计高速缓存集间损耗矫平技术时,可重新考虑使用数据移动。

[0033] 实现非易失性存储器高速缓存的集合地址重映射的另一选项是执行数据无效化。可以使用高速缓存线无效化,因为高速缓存中的数据可以在之后从较低级存储器再次回读。高速缓存的该特殊的特征为设计高速缓存集间损耗矫平技术提供了新的机会。

[0034] 相比于数据移动,无效化并不招致任何面积开销。因此,本公开的一方面修改了先前的主存储器损耗矫平技术,并且通过使用交换移位损耗矫平方案以及使用无效化来减少非易失性存储器高速缓存中的集间写差异的方式来增强这些技术。

[0035] 与用于非易失性主存储器的现有损耗矫平技术形成对比的是,交换移位方案是为非易失性存储器高速缓存所设计的。交换移位方案在改变集合地址映射时使用数据无效化而不是数据移动,从而减小面积和性能开销二者。

[0036] 交换移位方案的一个配置将对高速缓存物理集的映射移位以在集合之间轮转所存储的数据。然而,一次性将所有高速缓存集移位导致显著的性能开销。为了解决这一问题,高速缓存控制器 140 的交换移位方案一次仅交换对两个集合的映射,并且在完整的交换轮转之后,所有的高速缓存集可以用一步来被移位。

[0037] 在该配置中,高速缓存控制器 140 包括在交换移位方案中用以存储向该高速缓存所作的存储器写操作的数目的全局计数器 142,其由变量名“numWrite”指代。高速缓存控

制器 140 还包括用于存储当前交换值的交换寄存器 144 (SwapReg)。SwapReg 最初被设置为 0, 并且循环地从 0 改变到 N-1, 其中 N 是该高速缓存中集合的数目。高速缓存控制器 140 进一步包括存储当前移位值的移位寄存器 146 (ShiftReg)。ShiftReg 循环地从 0 改变到 N。这两个值 (SwapReg 和 ShiftReg) 由高速缓存控制器 140 使用以在移位交换方案中控制两个类型的轮转 (交换轮转和移位轮转)。

[0038] 首先描述交换轮转。当 numWrite 等于特定的预定义阈值 (“阈值”) 时, SwapReg 被递增 1, 并且当 SwapReg 被移动了 N-1 步时发生一个交换轮转。因此, 一个交换轮转由 N-1 个交换构成。

[0039] 其次描述移位轮转。在每个交换轮转后, ShiftReg 被递增 1, 并且当 ShiftReg 被移动了 N 步时, 发生一个移位轮转。因此, 每个移位轮转由 N 个交换轮转构成。

[0040] 图 2 是根据本公开的一方面的示出用于损耗矫平的示例轮转的示图 200。示图 200 包括高速缓存结构 202 以及一组轮转框 204、206、208、210、212、214、216 和 218。高速缓存结构 202 是包括高速缓存块 114、高速缓存路 108 或高速缓存集 116 的任何结构, 并且可以包括整个高速缓存存储器 100 本身, 尽管为了简化起见, 高速缓存结构 202 被表示为具有一群高速缓存集 202a、202b、202c 和 202d。第一高速缓存集 202a、第二高速缓存集 202b、第三高速缓存集 202c 和第四高速缓存集 202d 是为示图 200 的轮转示例示出的示例高速缓存集, 但是高速缓存集或者高速缓存结构 202 的子结构的数目并不限于四个。现在描述轮转过程。

[0041] 在轮转框 204 中, 高速缓存结构 202 处在其最初位置, 其中第一高速缓存集 202a 用 “0” 标记, 第二高速缓存集 202b 用 “1” 标记, 第三高速缓存集 202c 用 “2” 标记, 而第四高速缓存集用 “3” 标记。高速缓存集 202a、202b、202c 和 202d 的每一者中也可存储有数据。SwapReg 计数器对高速缓存集被交换了的次数进行计数。该计数器可以被初始化到 0。ShiftReg 计数器对整个高速缓存结构 202 被移位 (其中高速缓存结构 202 内的所有高速缓存集被移动了一个位置) 的次数进行计数, 其也可以被初始化并被设为 0。轮转的常规办法使用临时块并且随时间推移清除所有高速缓存集或高速缓存结构。然而, 对于示图 200 中所示的实现, 并未使用临时块, 高速缓存集的数据内容也未被删除或转储清除。这是真的, 因为数据并不是实际被交换了而是被保持在相同地点或位置上。仅仅是高速缓存集的位置被交换了, 并且仅仅是来自经交换的高速缓存集的内容被清除了, 而所有其他高速缓存集的内容保持完好。每个高速缓存集也被移动到它们的新位置并且新的数据被自动加载到新的高速缓存集中, 而不必移动任何数据。再一次地, 被交换的仅仅是高速缓存集的位置。所有事情都是用交换操作来做到的, 而非实际地移动数据。

[0042] 在轮转框 206, 第一高速缓存集 202a(0) 的位置被与第二高速缓存集 202b(1) 的位置作了交换。结果, 第二高速缓存集 202b(1) 现在是最首位或顶上的高速缓存集, 并且第一高速缓存集 202a(0) 变为第二高速缓存集, 其仅次于或低于第一高速缓存集 202a(0)。在一个实现中, 一旦该交换发生, 新数据将被自动重新加载到这些新高速缓存集 (一旦处于其新位置上) 中。在一个实现中, 经交换的高速缓存集将会保留其数据, 所以一旦它们被交换, 新数据将不会被自动重新加载到高速缓存集中。SwapReg 计数器 (其对高速缓存集被交换了的次数进行计数) 也可被递增 1。在轮转框 208 中, 第一高速缓存集 202a(0) 的位置 (现在在第二位置, 即第二高速缓存集 202b(1) 的老位置上) 被与第三高速缓存集 202c(2)

的位置作了交换。现在,第三高速缓存集 202c(2) 变为第二位置(第二高速缓存集 202b(1) 的先前位置),并且第一高速缓存集 202a(0) 变为第三位置(第三高速缓存集 202c(2) 的先前位置)。SwapReg 计数器接着可以再次被递增 1。

[0043] 在轮转框 210 中,第一高速缓存集 202a(0) 的位置(现在在第三位置,即第三高速缓存集 202c(2) 的老位置上)被与第四高速缓存集 202d(3) 的位置作了交换。现在,第四高速缓存集 202d(3) 变为第三位置(第三高速缓存集 202c(2) 的先前位置),并且第一高速缓存集 202a(0) 变为第四位置(第四高速缓存集 202d(3) 的先前位置)。SwapReg 计数器接着可以被初始化,因为所有高速缓存集已经被移位,并且 ShiftReg 计数器被递增,因为在三轮交换后,高速缓存结构 202 内的所有寄存器已被移位了一个位置。

[0044] 继续行至轮转框 212,第二高速缓存集 202b(1) 的位置被与第三高速缓存集 202c(2) 的位置作了交换。现在,第二高速缓存集 202(b)(1) 变为第二位置(第三高速缓存集 202c(2) 的先前位置),并且第三高速缓存集 202c(2) 变为第一位置(第二高速缓存集 202b(1) 的先前位置)。SwapReg 计数器也可以被递增 1。在轮转框 214,第二高速缓存集 202b(1) 的位置被与第四高速缓存集 202d(3) 的位置作了交换。现在,第四高速缓存集 202d(3) 变为第二位置(第二高速缓存集 202b(1) 的先前位置),并且第四高速缓存集 202d(3) 变为第二位置(第二高速缓存集 202b(1) 的先前位置)。SwapReg 计数器也被递增 1。在轮转框 216 中,第二高速缓存集 202(b)(1) 的位置被与第一高速缓存集 202(a)(0) 的位置作了交换。现在,第二高速缓存集 202(b)(1) 变为第四位置(第一高速缓存集 202a(0) 的先前位置),并且第一高速缓存集 202a(0) 变为第三位置(第二高速缓存集 202b(1) 的先前位置)。SwapReg 计数器接着被初始化并且 ShiftReg 计数器递增 1,因为在另三轮的交换之后,高速缓存结构 202 内的所有寄存器移动了两个位置。

[0045] 在一个实现中,整个高速缓存的内容没有被转储清除,因为所有事情都是由交换操作来做到的。进一步,交换操作并不涉及任何性能降级。通过使用交换操作以执行交换,“上移位”操作和/或跟踪并未发生。在一个实现中,数据并未被交换,仅是位置交换了,并且数据接着被重加载到新的经交换位置中。在一个实现中,不被使用的位置上的数据简单地被丢弃,诸如无效高速缓存集。在一个实现中,对于一次交换,对应于 SwapReg 计数器位置的高速缓存集(例如,高速缓存寄存器 [SwapReg]) 的内容以及对应于 SwapReg 计数器加一的高速缓存集(例如,高速缓存寄存器 [SwapReg+1]) 的内容可以被丢弃。接着,它们的物理位置被交换。在 N-1 次交换后(其中 N 是高速缓存结构中的高速缓存集的数目),该高速缓存结构内的所有物理位置(每个高速缓存集)全被移位了 1。

[0046] 在一个实现中,如图 1 中所示,SwapReg 和 ShiftReg 计数器可以被实现为高速缓存控制器 140 的交换寄存器 144 和移位寄存器 146 以用于存储数据。在这一情形中,这些寄存器被实现为对它们各自相应的计数器的值保持跟踪的某种形式的数值数据。

[0047] 图 3 是示出根据本公开的一方面的用于对高速缓存存储器进行损耗矫平的高速缓存控制器的操作的逻辑流程图 300。图 3 示出了在损耗矫平系统中 ShiftReg 和 SwapReg 计数器如何关于写活动来被更新。在框 302,确定是否已有高速缓存写。若有高速缓存写,那么在框 304,称为“numWrite”的高速缓存写计数器被递增一。若没有高速缓存写,那么该过程回到框 302 之前。numWrite 计数器还触发何时进行交换。还有某个预定的阈值,其决定多少个写足以发起高速缓存集的交换。例如,在框 306,确定 numWrite 计数器是否等于预定

阈值。若是,那么在框 308 中,发起交换,其将 numWrite 计数器更新至零值并且将 SwapReg 计数器更新至  $(\text{SwapReg}+1) \bmod (N-1)$  的值。Mod 是取模运算符并且 N 是所选高速缓存结构中的高速缓存集的总数。在框 306,若 numWrite 计数器并不等于该预定阈值,那么该过程回到框 304。在框 310,确定 SwapReg 计数器是否等于零。若是,那么在框 312,ShiftReg 计数器被更新到  $(\text{ShiftReg}+1) \bmod N$  的值。这将整个高速缓存结构移位一。在框 310,若 SwapReg 计数器并不等于零,那么该过程回到框 308。

[0048] 图 3 也可以用如下伪代码来表达:

```
[0049] If(有高速缓存写),
[0050] then:numWrite++;
[0051] If(numWrite == threshold),
[0052] then:numWrite = 0;
[0053] then:SwapReg = (SwapReg+1) mod (N-1);
[0054] if(SwapReg == 0),
[0055] then:ShiftReg = (ShiftReg+1) mod (N).
```

[0056] 图 4 是示出根据本公开的一方面的用于对高速缓存存储器进行损耗矫平的高速缓存控制器的操作的逻辑流程图 400。图 4 示出了存储表示写次数的 numWrite 计数器的全局计数器 404,每一次写由高速缓存写动作 402 表示。SwapReg 计数器(反映为“SwapReg”)以及 ShiftReg 计数器(表达为“ShiftReg”)可以被实现为用于分别存储 SwapReg 和 ShiftReg 值的数据寄存器。

[0057] 如关于图 3 所描述的,若有高速缓存写 402,那么全局计数器 404 被增加。若在 406 全局计数器等于预定阈值,则在框 408 发生交换,其将 SwapReg 计数器更新到  $(\text{SwapReg}+1) \bmod (N-1)$  的值,其中 mod 是取模运算符,并且 N 是总高速缓存集的数目。若 SwapReg 为零(410),那么在框 412,ShiftReg 计数器被更新到  $(\text{ShiftReg}+1) \bmod N$  的值。

[0058] 当逻辑集 (LS) 号传入以作为逻辑集号输入 414 时,物理集 (PS) 号可以基于三个不同情境而被计算为物理集号输出 418。

[0059] 首先,如逻辑框 416 中所示,若逻辑集号输入 414 等于 SwapReg 值,则这意味着该逻辑集合恰好是在该次轮转中应当被交换的那个高速缓存集。因此,该物理集被映射到 ShiftReg 的当前移位值,并且被输出以作为物理集号输出 418。

[0060] 其次,如同在逻辑框 416 中示出的,若逻辑集号输入 414 大于 SwapReg 值,则这意味着该高速缓存集在本次轮转中尚未被移位,并且保持与上一次轮转相同的映射。因此,物理集被映射到  $\text{LS}+\text{ShiftReg}$ 。在一个实现中,该映射是通过取  $(\text{LS}+\text{ShiftReg})$  值、并且执行与  $N$ (高速缓存集的数目)的取模运算、并且接着将结果所得的值赋给物理集来完成的。

[0061] 第三,如同在逻辑框 416 中示出的,若逻辑集号输入 414 小于 SwapReg 值(否则子句),则这意味着该高速缓存集在本次轮转中已被移位。因此,该物理集被映射到  $\text{LS}+\text{ShiftReg}+1$ 。在一个实现中,该映射是通过取  $(\text{LS}+\text{ShiftReg}+1)$  值、并且执行与  $N$ (高速缓存集的数目)的取模运算、并且接着将结果所得的值赋给物理集来完成的。

[0062] 类似于以上所讨论的操作的三个操作在逻辑框 422 中发生,但是是对于物理集输入号 424 发生的,从而计算逻辑集号输出 420。

[0063] 当高速缓存线需要被写回到较低级存储器时,逻辑集地址就被重新生成。从物理

集到逻辑集的映射是对称的。该映射方案也可以如图 2 中所见地被验证。因为 SwapReg 和 ShiftReg 随着增加的写计数而被改变,所以逻辑集与物理集之间的映射总是在改变,这确保了向不同物理集的写入是平衡的,从而减小了写差异。

[0064] 相比于常规高速缓存架构,移位交换损耗矫平方案中的集合索引转译仅仅增加了简单的算术运算并且可以被合并到行解码器中。此外,该一个循环的等待时间开销仅在较高级高速缓存访问较低级高速缓存未命中时被付出。

[0065] 图 5 是解说根据本公开的一方面的用于损耗矫平的方法的过程流程图 500。在框 502 中,由全局计数器通过针对每次存储器写操作将该全局计数器递增一来对存储器写操作的数目进行计数。在框 504,当全局计数器等于预定阈值时,交换操作被执行并且交换计数器被递增一。同样,执行交换操作包括重映射经交换的高速缓存集,这进而包括将这两个经交换高速缓存集内的脏数据置于写回缓冲器中,并且将这两个经交换高速缓存集中的其他数据无效化。

[0066] 在框 506,当交换计数器等于  $N-1$  时,移位计数器被递增一并且交换计数器被设回零。 $N$  是高速缓存存储器中的高速缓存集的数目。在框 508,当移位计数器等于  $N$  时,移位计数器被设回零。在框 510,输入高速缓存集号被转换成输出高速缓存集号。该输入或输出高速缓存集号可以是逻辑集号或物理集号。

[0067] 图 6 是解说根据本公开的一方面的用于损耗矫平的方法的过程流程图 600。在框 610,确定向高速缓存存储器的存储器写操作的数目(例如,numWrites)是否已达到了预定阈值。若是,则在框 612,对两个高速缓存集执行存储器位置交换操作。在框 614,来自这两个经交换高速缓存集之一的内容被清除。在框 616,经交换的这两个高速缓存集中的另一个的存储器内容被保持完好。在框 618,高速缓存集的交换被跟踪。在框 620,逻辑高速缓存集号被转换成物理高速缓存集号。若写操作的数目尚未达到该阈值,那么该过程停留在框 610 中。

[0068] 在交换移位损耗矫平方案中,集间写差异减小与在实验时间期间的移位轮转的数目有关。假定高速缓存中有  $N$  个集合,则一个移位轮转包括  $N$  个交换轮转,并且交换移位方案中的一个交换轮转需要  $N-1$  个交换。在每个移位轮转之后,所有高速缓存集被移位了  $N$  步,并且逻辑集索引被映射到它们原始的位置。因此,高速缓存被移位越多轮,写访问就被越平均地分布到每个高速缓存集。

[0069] 根据本公开的进一步方面,描述了用于高速缓存存储器的损耗矫平的高速缓存控制器。该高速缓存控制器包括用于通过对高速缓存存储器的存储器写操作的数目达到阈值时对高速缓存集执行多个存储器位置交换操作的方式来动态地轮转高速缓存存储器的高速缓存集的装置。每个交换操作可以包括仅清除来自经交换高速缓存集的内容,而保持其他高速缓存集的存储器内容完好。该动态轮转装置可以是高速缓存控制器 140。该设备进一步包括用于跟踪经交换高速缓存集以将逻辑高速缓存集号转换成物理高速缓存集号的装置。该跟踪装置可以是高速缓存控制器 140、全局计数器 142、交换寄存器 144、和 / 或移位寄存器 146。在另一方面,前述装置可以是被配置成执行由前述装置所陈述的功能的任何模块或任何设备。

[0070] 图 7 是示出其中可有利地采用本公开的一方面的示例性无线通信系统 700 的框图。出于解说目的,图 7 示出了三个远程单元 720、730 和 750 以及两个基站 740。将认识

到,无线通信系统可具有远多于此的远程单元和基站。远程单元 720、730 和 750 包括 IC 器件 725A、725C 和 725B,这些 IC 器件包括所公开的高速缓存存储器。将认识到,其他设备也可包括所公开的高速缓存存储器,诸如基站、交换设备、和网络装备。图 7 示出从基站 740 到远程单元 720、730 和 750 的前向链路信号 780,以及从远程单元 720、730 和 750 到基站 740 的反向链路信号 790。

[0071] 在图 7 中,远程单元 720 被示为移动电话,远程单元 730 被示为便携式计算机,而远程单元 750 被示为无线本地环路系统中的固定位置远程单元。例如,这些远程单元可以是移动电话、手持式个人通信系统 (PCS) 单元、便携式数据单元 (诸如个人数据助理)、启用全球定位系统 (GPS) 的设备、导航设备、机顶盒、音乐播放器、视频播放器、娱乐单元、固定位置数据单元 (诸如仪表读数装备)、或者存储或取回数据或计算机指令的其他设备、或者其组合。尽管图 7 解说了根据本公开的各方面的远程单元,但本公开并不被限定于所解说的这些示例性单元。本公开的各方面可以合适地在包括所公开的高速缓存存储器的许多设备中采用。

[0072] 图 8 是解说用于半导体组件 (诸如以上公开的高速缓存存储器) 的电路、布局以及逻辑设计的设计工作站的框图。设计工作站 800 包括硬盘 801,该硬盘 801 包含操作系统软件、支持文件、以及设计软件 (诸如 Cadence 或 OrCAD)。设计工作站 800 还包括促成对电路 810 或半导体组件 812 (诸如高速缓存存储器) 的设计的显示器 802。提供存储介质 804 以用于有形地存储电路设计 810 或半导体组件 812。电路设计 810 或半导体组件 812 可以文件格式 (诸如 GDSII 或 GERBER) 存储在存储介质 804 上。存储介质 804 可以是 CD-ROM、数字化通用碟 (DVD)、硬盘、闪存、或者其他恰适的设备。此外,设计工作站 800 包括用于从存储介质 804 接受输入或者将输出写入存储介质 804 的驱动装置 803。

[0073] 存储介质 804 上记录的数据可指定逻辑电路配置、用于光刻掩模的图案数据、或者用于串写工具 (诸如电子束光刻) 的掩模图案数据。该数据可进一步包括与逻辑仿真相关联的逻辑验证数据,诸如时序图或网电路。在存储介质 804 上提供数据通过减少用于设计半导体晶片的工艺数目来促成电路设计 810 或半导体组件 812 的设计。

[0074] 对于固件和 / 或软件实现,这些方法体系可以用执行本文所描述功能的模块 (例如,规程、函数等等) 来实现。有形地体现指令的机器可读介质可被用来实现本文所述的方法体系。例如,软件代码可被存储在存储器中并由处理器单元来执行。存储器可以在处理器单元内或在处理器单元外部实现。如本文所用的,术语“存储器”是指长期、短期、易失性、非易失性类型存储器、或其他存储器,而并不限于特定类型的存储器或存储器数目、或记忆存储在其上的介质的类型。

[0075] 如果以固件和 / 或软件实现,则功能可作为一条或多条指令或代码存储在计算机可读介质上。示例包括编码有数据结构的计算机可读介质和编码有计算机程序的计算机可读介质。计算机可读介质包括物理计算机存储介质。存储介质可以是能被计算机访问的可用介质。作为示例而非限定,此类计算机可读介质可包括 RAM、ROM、EEPROM、CD-ROM 或其他光盘存储、磁盘存储或其他磁存储设备、或能被用来存储指令或数据结构形式的期望程序代码且能被计算机访问的任何其他介质;如本文中所使用的盘 (disk) 和碟 (disc) 包括压缩碟 (CD)、激光碟、光碟、数字多用碟 (DVD)、软盘和蓝光碟,其中盘常常磁性地再现数据,而碟用激光光学地再现数据。上述的组合应当也被包括在计算机可读介质的范围内。

[0076] 除了存储在计算机可读介质上,指令和 / 或数据还可作为包括在通信装置中的传输介质上的信号来提供。例如,通信装置可包括具有指示指令和数据的信号的收发机。这些指令和数据被配置成使一个或多个处理器实现权利要求中叙述的功能。

[0077] 尽管已详细描述了本公开及其优势,但是应当理解,可在本文中作出各种改变、替代和变更而不会脱离如由所附权利要求所定义的本公开的技术。例如,诸如“上方”和“下方”之类的关系术语是关于基板或电子器件使用的。当然,如果该基板或电子器件被颠倒,则上方变成下方,反之亦然。此外,如果是侧面取向的,则上方和下方可指代基板或电子器件的侧面。而且,本申请的范围并非旨在被限定于说明书中所描述的过程、机器、制造、物质组成、装置、方法和步骤的特定配置。如本领域的普通技术人员将容易从本公开领会到的,根据本公开,可以利用现存或今后开发的与本文所描述的相应配置执行基本相同的功能或实现基本相同结果的过程、机器、制造、物质组成、装置、方法或步骤。因此,所附权利要求旨在将这样的过程、机器、制造、物质组成、装置、方法或步骤包括在其范围内。

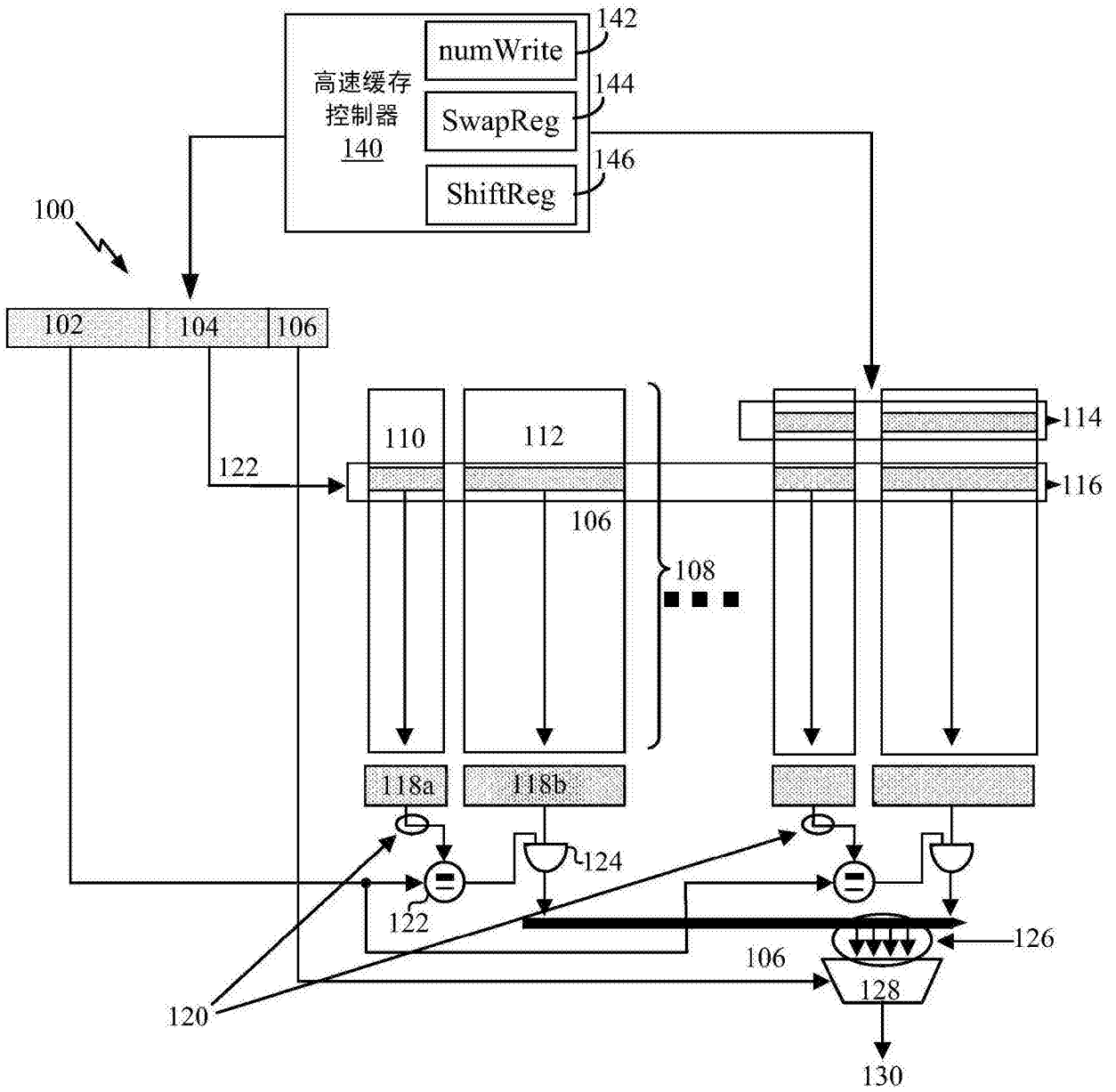


图 1

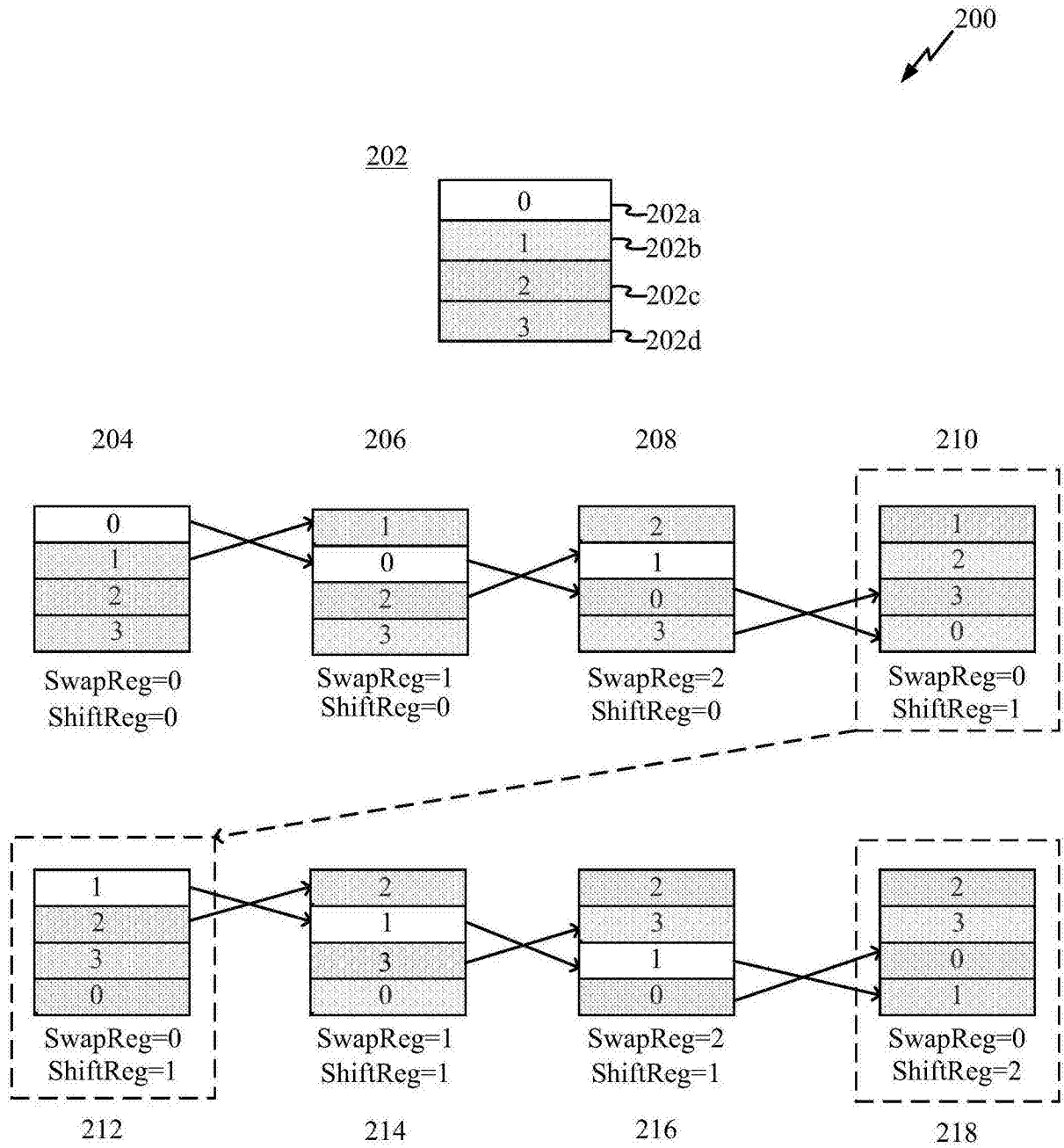


图 2

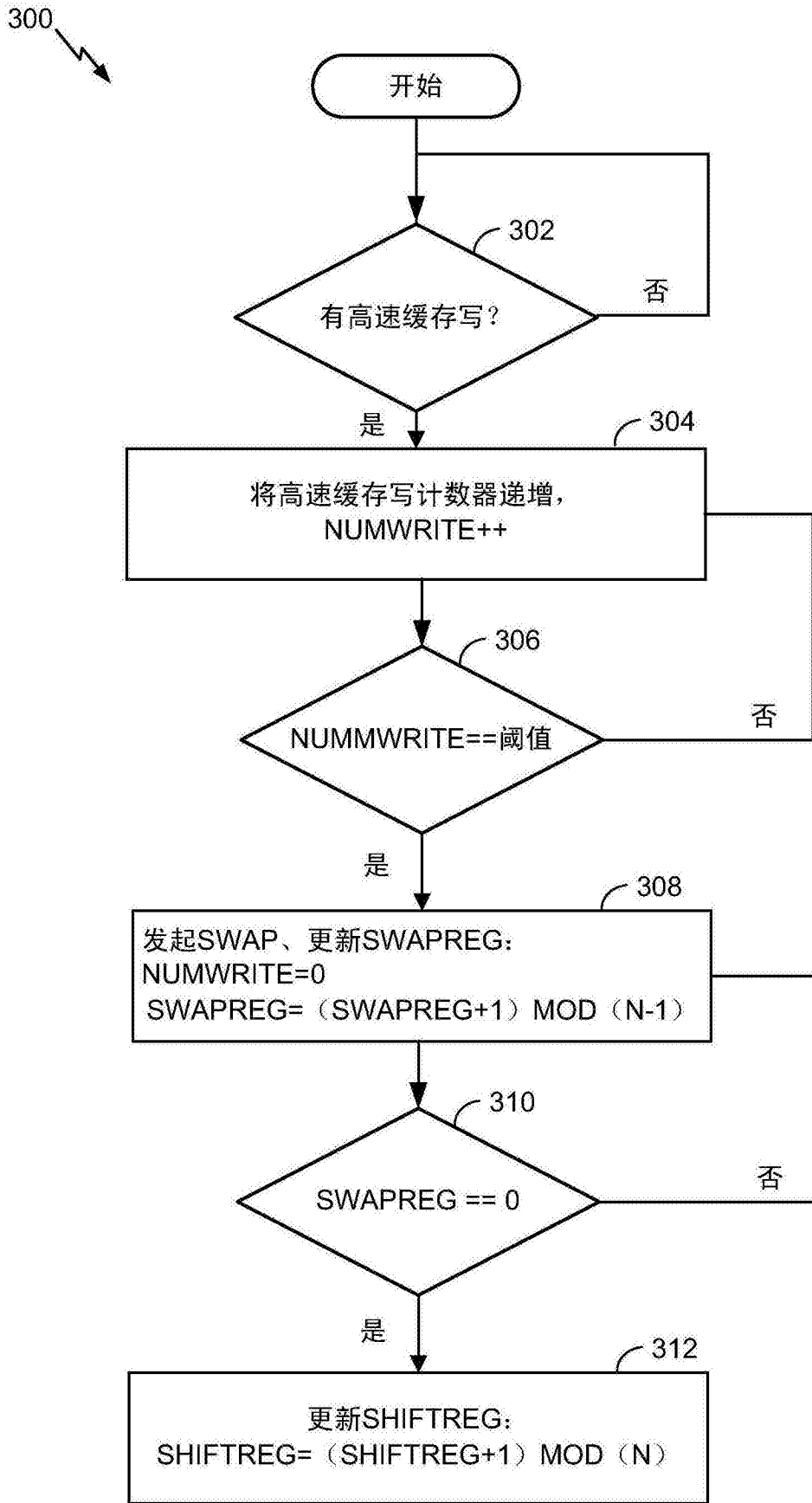


图 3

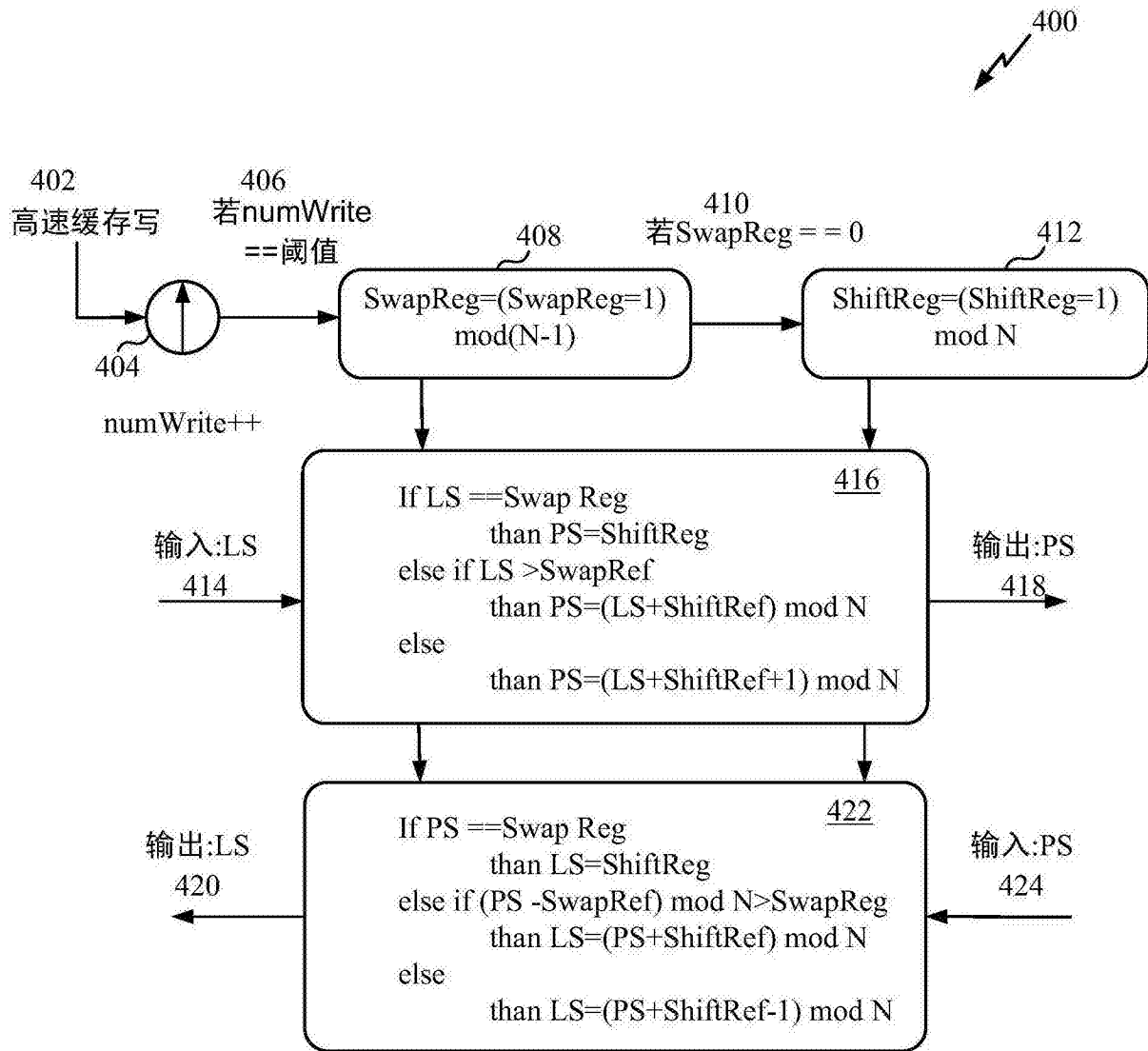


图 4

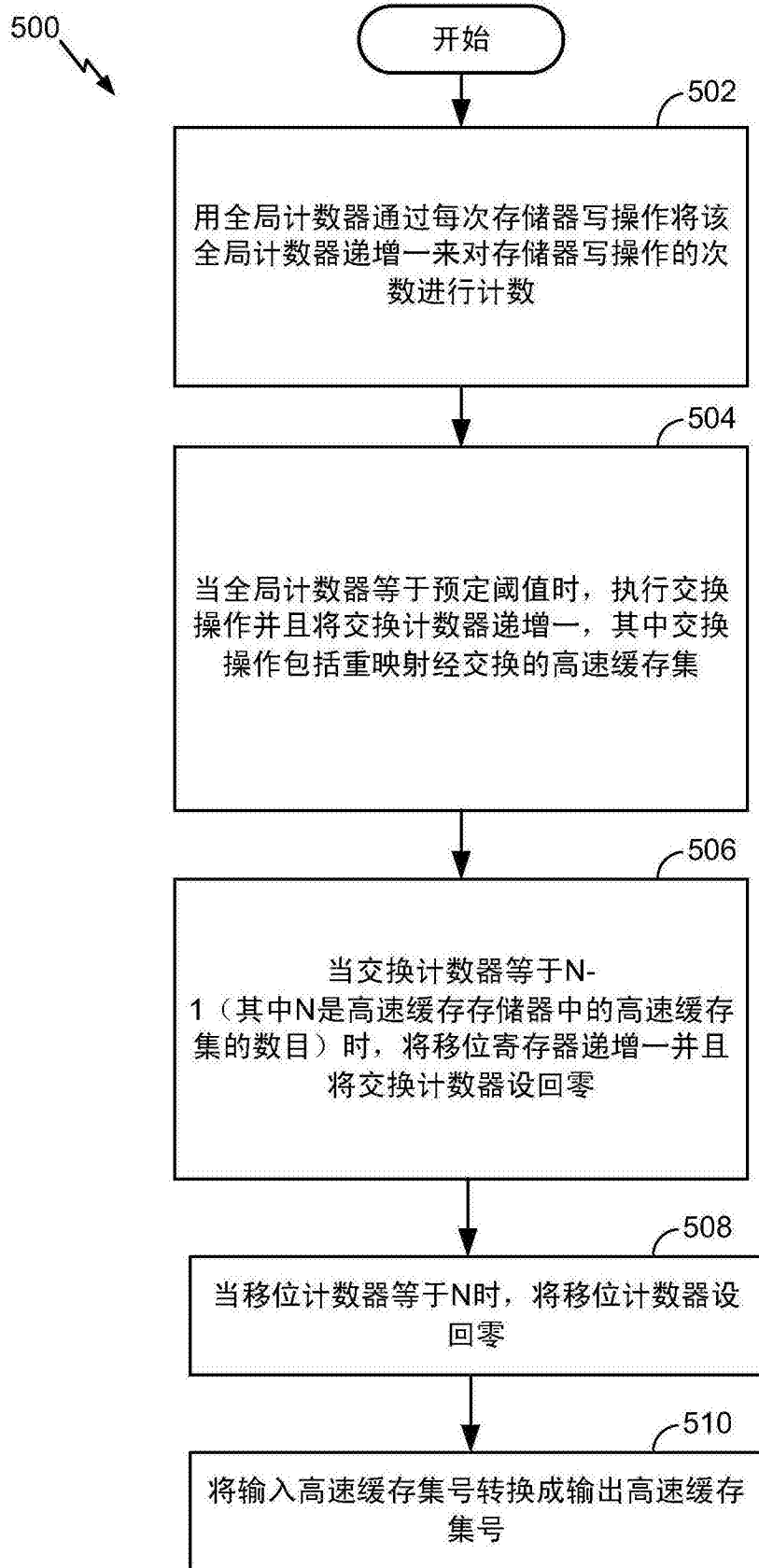


图 5

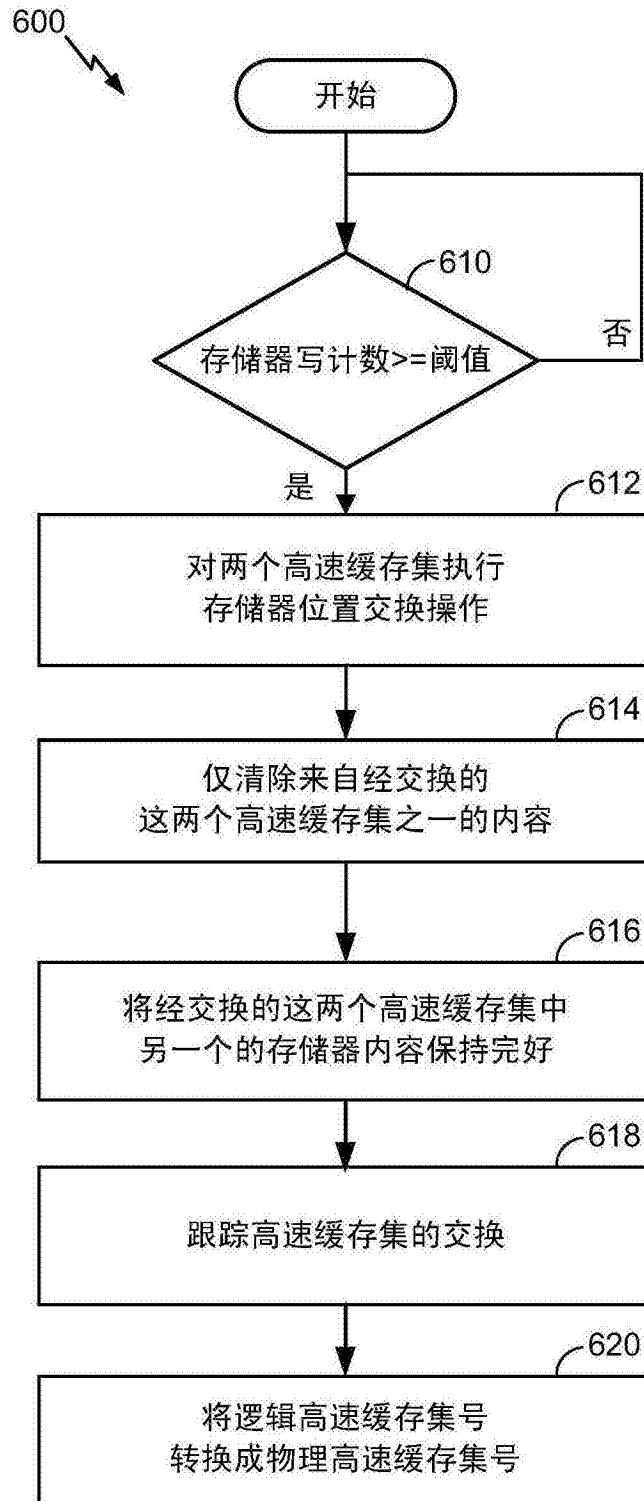


图 6

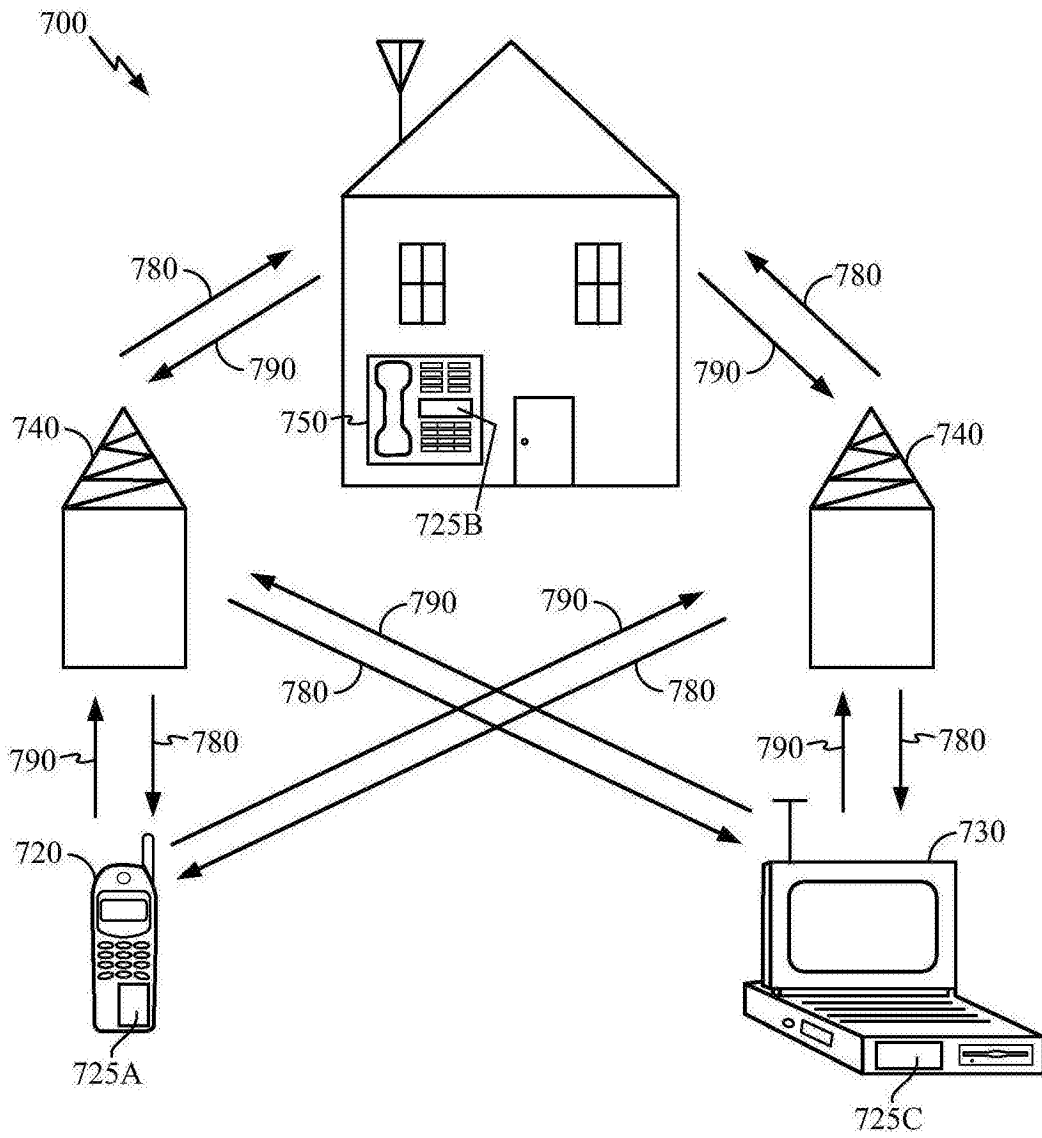


图 7

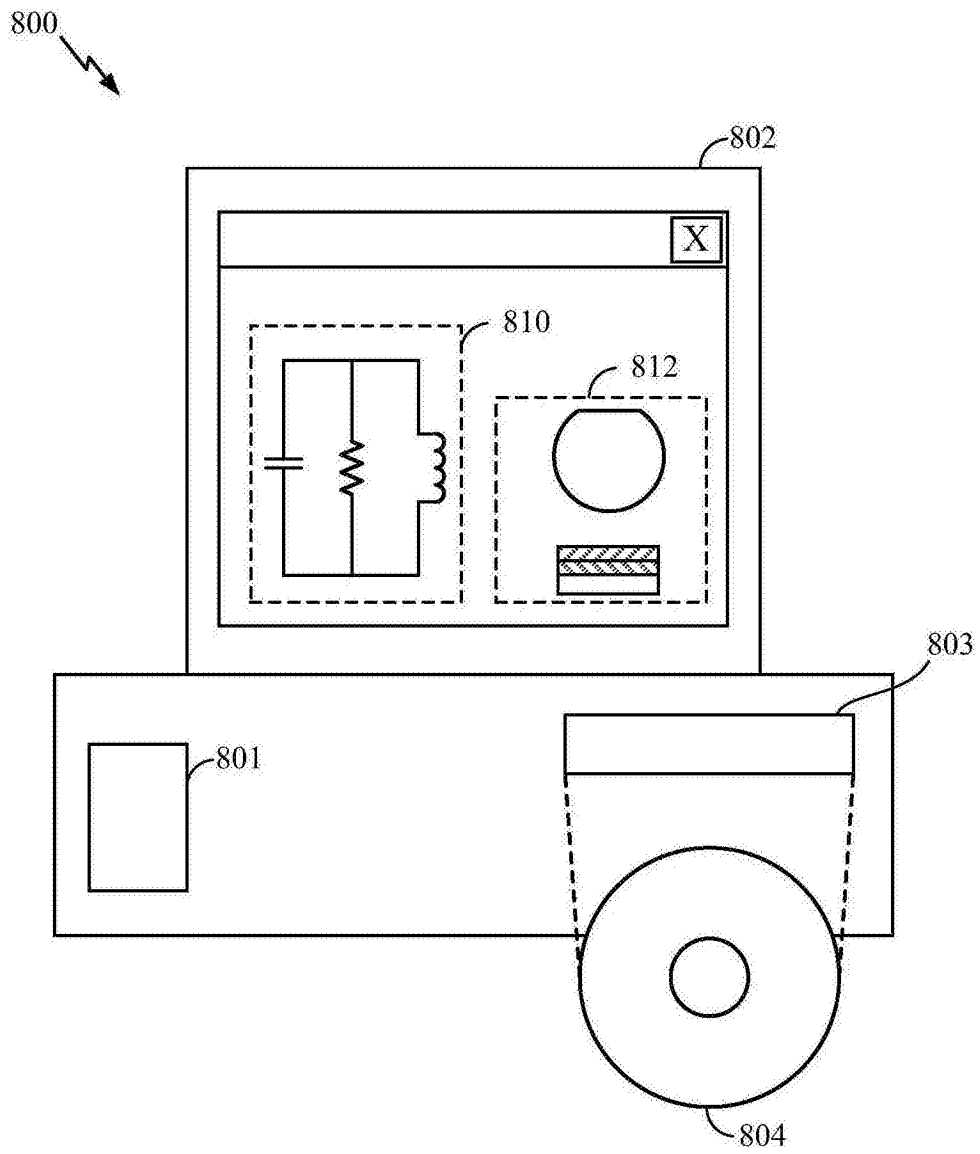


图 8