



(10) 授权公告号 CN 112424743 B

(45) 授权公告日 2024. 08. 02

(21) 申请号 201980046961.7

(22) 申请日 2019.07.24

(65) 同一申请的已公布的文献号
申请公布号 CN 112424743 A

(43) 申请公布日 2021.02.26

(30) 优先权数据
16/058,102 2018.08.08 US(85) PCT国际申请进入国家阶段日
2021.01.13(86) PCT国际申请的申请数据
PCT/US2019/043300 2019.07.24(87) PCT国际申请的公布数据
W02020/033149 EN 2020.02.13(73) 专利权人 美光科技公司
地址 美国爱达荷州

(72) 发明人 D·巴维什 T·C·迈耶罗维茨

(74) 专利代理机构 北京律盟知识产权代理有限公司
11287
专利代理师 王龙(51) Int.Cl.
G06F 3/06 (2006.01)
G06F 13/16 (2006.01)(56) 对比文件
US 8667254 B1, 2014.03.04
US 5561823 A, 1996.10.01
WO 2017192346 A1, 2017.11.09
US 2009175084 A1, 2009.07.09
US 2015288977 A1, 2015.10.08

审查员 夏玫

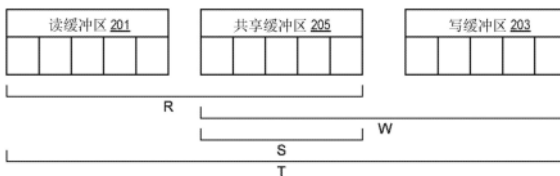
权利要求书4页 说明书18页 附图13页

(54) 发明名称

存储器系统中用于读请求和写请求的缓冲区管理

(57) 摘要

一种存储器系统,其具有存储器组件和处理装置,所述处理装置用于:从主机系统接收写请求;在所述写请求在所述存储器系统中待定的第一时间段期间,将所述写请求的数据存储在所述存储器系统的缓冲区的一部分中;从所述主机系统接收读请求;以及在所述读请求在所述存储器系统中待定的第二时间段期间,将所述读请求的数据存储在所述缓冲区的一部分中。所述缓冲区的存储所述写请求的所述数据的所述部分与所述缓冲区的存储所述读请求的所述数据的所述部分至少部分地重叠。



1. 一种存储器系统,其包括:

多个存储器组件;以及

主机接口,其经配置以经由包括命令总线、数据总线、事务总线和元数据总线的多个总线连接到主机系统,所述主机接口经由所述命令总线从所述主机系统接收从所述存储器组件读取数据的第一请求和向所述存储器组件写入数据的第二请求;

缓冲区,所述缓冲区经配置以在所述第一请求和所述第二请求的执行完成之前缓冲与所述第一请求和第二请求相关联的数据,其中所述缓冲区的至少一部分经配置以在不同时间段中缓冲用于不同类型的请求的数据,所述不同类型的请求包括读取数据的第一类型请求和写入数据的第二类型请求;

处理装置,其与所述多个存储器组件操作性地耦合且用于:

通过所述主机接口向所述主机系统传输初始指示,所述初始指示表示所述缓冲区中可用于存储与所述不同类型的请求相关联的数据的缓冲区空间的第一数量,以致使所述主机系统基于从所述主机系统传输的相应请求和在所述主机系统中接收的对所述相应请求的响应来计数所述存储器系统中可用于储存与所述不同类型的请求相关联的数据的缓冲区空间的数量;

经由所述命令总线从所述主机系统接收写请求;

在所述写请求在所述存储器系统中待定的第一时间段期间,将所述写请求的数据存储在所述存储器系统的所述缓冲区的一部分中;

经由所述命令总线从所述主机系统接收读请求;

在所述读请求在所述存储器系统中待定的第二时间段期间,将所述读请求的数据存储在所述缓冲区的一部分中;

保留所述缓冲区的第一部分,针对读请求和写请求共享所述第一部分;

在所述缓冲区的所述第一部分中缓冲写请求和读请求;

保留所述缓冲区的第二部分;

在所述第二部分中缓冲读请求但不缓冲写请求;

执行所述写请求;

执行所述读请求以从所述存储器组件检索第一数据;

经由所述事务总线向所述主机系统传输由所述读请求请求的所述第一数据可用的指示;

从所述主机系统接收用于发送所述第一数据的命令;以及

经由所述元数据总线并响应于发送所述第一数据的所述命令,传输作为至少所述写请求的完成的结果而释放的缓冲区空间的数量的指示,以致使所述主机系统确定所述缓冲区中可用于存储与所述不同类型的附加请求相关联的数据的缓冲区空间的第二数量;

其中,所述缓冲区的存储所述写请求的所述数据的所述部分与所述缓冲区的存储所述读请求的所述数据的所述部分至少部分地重叠;且

其中所述第一时间段与所述第二时间段不重叠。

2. 根据权利要求1所述的存储器系统,其中所述处理装置进一步用于:

保留所述缓冲区的第三部分;以及

在所述第三部分中缓冲写请求但不缓冲读请求。

3. 根据权利要求2所述的存储器系统,其中所述存储器组件包含非易失性存储器;并且所述处理装置将根据非易失性双列直插式存储器模块的通信协议从所述主机系统接收所述写请求和所述读请求。

4. 一种用于缓冲区管理的方法,其包括:

在主机系统中,从具有缓冲区并经由包括命令总线、数据总线、事务总线和元数据总线的多个总线耦合到所述主机系统的存储器系统接收表示所述缓冲区中可用于存储与不同类型的请求相关联的数据的缓冲区空间的量的初始指示;

在所述主机系统中存储标识可用于所述主机系统向所述存储器系统传输读请求和写请求的所述存储器系统的缓冲区容量的量,所述存储器系统具有多个存储器组件,且所述缓冲区用于缓冲与从所述存储器组件读取数据的第一类型请求和向所述存储器组件写入数据的第二类型请求相关联的数据,其中所述缓冲区的至少一部分经配置以可用于在不同时间段存储所述不同类型的请求,所述不同类型的请求包括所述第一类型请求和所述第二类型请求;

经由所述命令总线通过所述主机系统向所述存储器系统传输从所述存储器组件中检索第一数据的读请求;

通过所述主机系统基于所述读请求而减少所述量;

经由所述命令总线通过所述主机系统向所述存储器系统传输将数据存储在所述存储器组件中的写请求;

通过所述主机系统基于所述写请求而减少所述量;

在所述主机系统中存储用于从所述主机系统向所述存储器系统传输读请求的所述存储器系统的可用读缓冲区容量的量;

通过所述主机系统基于所述读请求而减少所述可用读缓冲区容量的量在所述主机系统中存储用于从所述主机系统向所述存储器系统传输写请求的所述存储器系统的可用写缓冲区容量的量;

通过所述主机系统基于所述写请求而减少所述可用写缓冲区容量的量;

在所述存储器系统执行所述读请求和所述写请求之后,在所述主机系统中经由所述事务总线从所述存储器系统接收由所述读请求请求的所述第一数据可用的指示;

从所述主机系统向所述存储器系统传输请求所述存储器系统发送所述第一数据的命令;以及

响应于所述命令,

在所述主机系统中经由所述数据总线从所述存储器系统接收由所述读请求请求的所述第一数据;

在所述主机系统中,经由所述元数据总线从所述存储器系统接收由于至少所述写请求的完成而释放的缓冲区容量的量的指示;

由所述主机系统根据由于至少所述写请求的所述完成而释放的缓冲区容量的所述量来增加所述写缓冲区容量的所述量;以及

由所述主机系统基于所述接收由所述读请求请求的所述第一数据来增加所述可用读缓冲区容量的所述量。

5. 根据权利要求4所述的方法,其进一步包括:

在所述主机系统中从所述存储器系统接收标识完成一或多个写请求之后变得可用的所述存储器系统的缓冲区容量的所述量的数据;以及

根据标识完成所述一或多个写请求之后变得可用的所述存储器系统的所述缓冲区容量的量的所述数据,通过所述主机系统增加标识可用于从所述主机系统向所述存储器系统传输读请求和写请求的所述存储器系统的缓冲区容量的所述量。

6. 一种非暂时性计算机可读存储媒体,其存储指令,所述指令当由处理装置执行时使所述处理装置:

在主机系统中,从具有缓冲区并经由包括命令总线、数据总线、事务总线和元数据总线的多个总线耦合到所述主机系统的存储器系统接收表示所述缓冲区中可用于存储与不同类型的请求相关联的数据的缓冲区空间的量的初始指示;

在所述主机系统中存储具有多个存储器组件的所述存储器系统的读缓冲区容量的量,且所述缓冲区用于缓冲与从所述存储器组件读取数据的第一类型请求和向所述存储器组件写入数据的第二类型请求相关联的数据,其中所述缓冲区的至少一部分经配置以可用于在不同时间段存储所述不同类型的请求,所述不同类型的请求包括所述第一类型请求和所述第二类型请求;

经由所述命令总线通过所述主机系统向所述存储器系统传输从所述存储器组件中检索第一数据的读请求;

通过所述主机系统基于所述读请求而减少所述读缓冲区容量的量;

在所述主机系统中存储所述存储器系统的总缓冲区容量的量;

通过所述主机系统基于所述读请求而减少所述总缓冲区容量的量;

经由所述命令总线通过所述主机系统向所述存储器系统传输将数据存储在所述存储器组件中的写请求;

通过所述主机系统基于所述写请求而减少所述总缓冲区容量的量;

在所述主机系统中存储所述存储器系统的写缓冲区容量的量;

通过所述主机系统基于所述读请求而减少所述写缓冲区容量的所述量;

在所述存储器系统执行所述读请求和所述写请求之后,在所述主机系统中,经由所述事务总线从所述存储器系统接收由所述读请求请求的所述第一数据可用的指示;

从所述主机系统向所述存储器系统传输请求所述存储器系统发送所述第一数据的命令;以及

响应于所述命令,

在所述主机系统中经由所述数据总线从所述存储器系统接收由所述读请求请求的所述第一数据;

在所述主机系统中,经由所述元数据总线从所述存储器系统接收由于至少所述写请求的完成而释放的缓冲区容量的量的指示;

由所述主机系统根据由于至少所述写请求的所述完成而释放的缓冲区容量的所述量来增加所述写缓冲区容量的所述量;以及

由所述主机系统基于接收由所述读请求请求的所述第一数据来增加所述读缓冲区容量的所述量。

7. 根据权利要求6所述的非暂时性计算机可读存储媒体,其中所述指令当由所述处理

装置执行时使所述处理装置：

在所述主机系统中从所述存储器系统接收标识完成一或多个写请求之后变得可用的缓冲区容量的所述量的数据；以及

基于标识所述存储器系统中所述完成所述一或多个写请求之后变得可用的所述缓冲区容量的量的所述数据，通过所述主机系统增加所述写缓冲区容量的量。

8. 根据权利要求7所述的非暂时性计算机可读存储媒体，其中所述指令当由所述处理装置执行时使所述处理装置：

基于标识所述完成所述一或多个写请求之后变得可用的所述缓冲区容量的量的所述数据，通过所述主机系统增加所述总缓冲区容量的量。

9. 根据权利要求6所述的非暂时性计算机可读存储媒体，其中所述指令当由所述处理装置执行时使所述处理装置：

在所述主机系统中存储标识用于缓冲读请求和写请求的所述存储器系统的缓冲区容量的共享部分的数据。

10. 根据权利要求9所述的非暂时性计算机可读存储媒体，其中所述指令当由所述处理装置执行时使所述处理装置：

基于标识用于缓冲读请求和写请求的所述缓冲区容量的共享部分的所述数据，计算可用于传输读请求和写请求的总缓冲区容量的量。

11. 根据权利要求9所述的非暂时性计算机可读存储媒体，其中所述指令当由所述处理装置执行时使所述处理装置：

基于标识用于缓冲读请求和写请求的缓冲区容量的共享部分的所述数据，计算可用于传输读请求的读缓冲区容量的量。

存储器系统中用于读请求和写请求的缓冲区管理

[0001] 相关申请

[0002] 本申请要求于2018年8月8日提交的名称为“存储器系统中用于读请求和写请求的缓冲区管理(Buffer Management in Memory Systems for Read and Write Requests)”的第16/058,102号美国专利申请的权益,所述美国专利申请的全部公开内容以引用的方式并入本文。

技术领域

[0003] 本文公开的至少一些实施例总体上涉及存储器系统,且更具体地,涉及存储器系统中用于读请求和写请求的缓冲区管理。

背景技术

[0004] 存储器子系统可以是存储系统,例如固态驱动器(SSD);其也可以是存储器模块,例如非易失性双列直插式存储器模块(NVDIMM),并且可以包含存储数据的一或多个存储器组件。存储器组件可以是例如非易失性存储器组件和易失性存储器组件。一般来说,主机系统可以利用存储器子系统以在存储器组件处存储数据且从存储器组件检索数据。

[0005] 标准化的通信协议允许主机系统与存储器子系统进行通信,以存储数据和检索数据。

[0006] 例如,JEDEC(电子装置工程设计联合会)固态技术协会提出了一种用于主机系统与NVDIMM-P存储器模块之间的通信的“DDR5 NVDIMM-P总线协议”。JEDEC委员会信函投票委员会:JC-45.6,委员会项目编号2261.13D,主题:“拟议的DDR5 NVDIMM-P总线协议”对此协议进行了详细描述,其在此以引用的方式整体并入本文中。

附图说明

[0007] 根据以下给出的详细描述和本公开的各种实施例的附图,将更充分地理解本公开。

[0008] 图1示出根据本公开的一些实施例的具有存储器子系统的实例计算系统。

[0009] 图2示出根据本公开的一些实施例的包含请求管理器和缓冲区管理器的实例计算系统。

[0010] 图3A-3C示出根据本公开的一些实施例的实例缓冲区配置。

[0011] 图4-9示出的序列图示出根据本公开的一些实施例的结合读操作或写操作的缓冲区容量跟踪的实例。

[0012] 图10是根据本公开的一些实施例的用于管理读请求和写请求共享的缓冲区的实例方法的流程图。

[0013] 图11和12是根据本公开的一些实施例的用于管理对存储器子系统的请求的实例方法的流程图。

[0014] 图13是本公开的实施例可在其中操作的实例计算机系统的框图。

具体实施方式

[0015] 本公开的至少一些方面针对改进的缓冲区管理技术,其可以减少在缓冲从主机系统传输到存储器子系统以供执行的读请求和写请求时的通信瓶颈。存储器子系统在下文也称为“存储器装置”。存储器子系统的实例是经由存储器总线连接到中央处理单元(CPU)的存储器模块,例如双列直插式存储器模块(DIMM)、小尺寸DIMM(SO-DIMM)、非易失性双列直插式存储器模块(NVDIMM)等。存储器子系统的另一实例是存储系统,例如固态驱动器(SSD)。在一些实施例中,存储器子系统是提供存储器功能和存储功能两者的混合式存储器/存储子系统。通常,主机系统可以利用包含一或多个存储器组件的存储器子系统。主机系统可以提供要存储在存储器子系统处的数据,并且可以请求从存储器子系统中检索数据。

[0016] 在一些计算机系统(例如使用NVDIMM-P总线连接的主机系统和存储器子系统)中,子系统可以缓冲来自主机系统的读命令和写命令,以在并非预定的时间段内执行。不需要存储器子系统按照在存储器子系统中接收命令的顺序来执行读命令和写命令。然而,当存储器子系统在读缓冲区已满时,将阻止主机系统向存储器子系统发送进一步的读命令,直到读缓冲区中的某些读命令已完成并从读缓冲区中删除为止;并且当存储器子系统在写缓冲区已满时,将阻止主机系统向存储器子系统发送进一步的写命令,直到写缓冲区中的某些写命令已完成并从写缓冲区中删除为止。

[0017] 本公开的至少一些方面通过允许使用存储器子系统的缓冲区的至少一部分来根据需要来缓冲读命令和写命令来解决上述和其它缺陷。当主机系统需要发送更多写命令时,在读写之间共享的这一部分可用于将写命令传输到存储器子系统。当主机系统需要发送更多读命令时,在读写之间共享的这一部分可用于将读命令传输到存储器子系统。使用针对读命令的可用缓冲区容量的标识、针对写命令的可用缓冲区容量的标识和/或针对读命令和写命令的可用总缓冲区容量的标识,主机系统可以跟踪针对读命令和写命令的可用缓冲区容量命令。

[0018] 图1示出根据本公开的一些实施例的具有存储器子系统110的实例计算系统100。存储器子系统110可以包含媒体,例如存储器组件109A到109N。存储器组件109A到109N可以是易失性存储器组件、非易失性存储器组件或此类组件的组合。在一些实施例中,存储器子系统110是存储器模块。存储器模块的实例包含DIMM、NVDIMM和NVDIMM-P。在一些实施例中,存储器子系统是存储系统。存储系统的实例是SSD。在一些实施例中,存储器子系统110是混合存储器/存储子系统。一般来说,计算环境可以包含使用存储器子系统110的主机系统120。例如,主机系统120可以将数据写入到存储器子系统110以及从存储器子系统110中读取数据。

[0019] 主机系统120可以是计算装置,例如台式计算机、膝上型计算机、网络服务器、移动装置,或包含存储器和处理装置的此类计算装置。主机系统120可以包含或耦合到存储器子系统110,使得主机系统120可以从存储器子系统110中读取数据或将数据写入到所述存储器子系统。主机系统120可以经由物理主机接口耦合到存储器子系统110。如本文所使用,“耦合到”通常是指组件之间的连接,其可以是间接通信连接或直接通信连接(例如,没有中间组件),无论是无线还是有线的,包含例如电连接、光学连接、磁连接等的连接。物理主机接口的实例包含但不限于串行高级技术附件(SATA)接口、外围组件互连快速(PCIe)接口、

通用串行总线 (USB) 接口、光纤通道、串行连接的 SCSI (SAS)、双倍数据速率 (DDR) 存储器总线等。物理主机接口可用于在主机系统120与存储器子系统110之间传输数据。当存储器子系统110通过PCIe接口与主机系统120耦合时,主机系统120可以进一步利用NVM快速 (NVMe) 接口访问存储器组件109A到109N。物理主机接口可以提供用于在存储器子系统110与主机系统120之间传递控制、地址、数据和其它信号的接口。图1示出存储器子系统110作为实例。一般来说,主机系统120可经由同一通信连接、多个单独通信连接和/或通信连接的组合来访问多个存储器子系统。

[0020] 主机系统120包含处理装置118和控制器116。主机系统120的处理装置118可以是例如微处理器、中央处理单元 (CPU)、处理器的处理核心、执行单元等。在一些情况下,控制器116可称为存储器控制器、存储器管理单元和/或起始器。在一个实例中,控制器116控制通过耦合在主机系统120与存储器子系统110之间的总线进行的通信。

[0021] 通常,控制器116可以向存储器子系统110发送期望访问存储器组件109A到109N的命令或请求。控制器116可进一步包含与存储器子系统110通信的接口电路系统。接口电路系统可以将从存储器子系统110接收的响应转换成主机系统120的信息。

[0022] 主机系统120的控制器116可以与存储器子系统110的控制器115通信以执行操作,例如在存储器组件109A到109N处读取数据、写入数据或擦除数据以及其它这类操作。在一些情况下,控制器116被集成在处理装置118的同一包装内。在其它情况下,控制器116与处理装置118分开包装。控制器116和/或处理装置118可以包含硬件,例如一或多个集成电路和/或离散组件、缓冲存储器、高速缓存存储器或其组合。控制器116和/或处理装置118可以是微控制器、专用逻辑电路系统 (例如现场可编程门阵列 (FPGA)、专用集成电路 (ASIC) 等) 或另一合适的处理器。

[0023] 存储器组件109A到109N可以包含不同类型的非易失性存储器组件和/或易失性存储器组件的任何组合。非易失性存储器组件的实例包含与非 (NAND) 类型快闪存储器。存储器组件109A到109N中的每一者可以包含一或多个存储器单元阵列,例如单级单元 (SLC) 或多级单元 (MLC) (例如,三级单元 (TLC) 或四级单元 (QLC))。在一些实施例中,特定存储器组件可以包含存储器单元的SLC部分和MLC部分两者。存储器单元中的每一者可以存储由主机系统120使用的一或多个数据位 (例如,数据块)。尽管描述了例如NAND类型快闪存储器的非易失性存储器组件,但存储器组件109A到109N可以基于任何其它类型的存储器,例如易失性存储器。在一些实施例中,存储器组件109A到109N可以是但不限于随机存取存储器 (RAM)、只读存储器 (ROM)、动态随机存取存储器 (DRAM)、同步动态随机存取存储器 (SDRAM)、相变存储器 (PCM)、磁性随机存取存储器 (MRAM)、自旋转移力矩 (STT) -MRAM、铁电随机存取存储器 (FeTRAM)、铁电RAM (FeRAM)、导电桥接RAM (CBRAM)、电阻式随机存取存储器 (RRAM)、基于氧化物的RRAM (OxRAM)、或非 (NOR) 快闪存储器、电可擦除可编程只读存储器 (EEPROM)、基于纳米线的非易失性存储器、并入有忆阻器技术的存储器,以及非易失性存储器单元的交叉点阵列。非易失性存储器的交叉点阵列可以结合可堆叠交叉网格化数据存取阵列基于体电阻的改变来进行位存储。另外,与许多基于快闪的存储器相比,交叉点非易失性存储器可以执行就地写操作,其中可以对非易失性存储器单元进行编程,而无需事先擦除非易失性存储器单元。此外,存储器组件109A到109N的存储器单元可分组为存储器页或数据块,其可指代用于存储数据的存储器组件的单元。

[0024] 存储器子系统110的控制器115可以与存储器组件109A到109N通信以执行操作,例如在存储器组件109A到109N处读取数据、写入数据或擦除数据以及其它这类操作(例如,响应于由控制器116在命令总线上调度的命令)。控制器115可以包含硬件,例如一或多个集成电路和/或离散组件、缓冲存储器或其组合。控制器115可以是微控制器、专用逻辑电路系统(例如现场可编程门阵列(FPGA)、专用集成电路(ASIC)等)或另一合适的处理器。控制器115可以包含经配置以执行存储在本地存储器119中的指令的处理装置117(处理器)。在所示的实例中,控制器115的本地存储器119包含经配置以存储指令的嵌入式存储器,所述指令用于执行控制存储器子系统110的操作(包含处理存储器子系统110与主机系统120之间的通信)的各种处理、操作、逻辑流以及例程。在一些实施例中,本地存储器119可以包含存储器寄存器,其存储存储器指针、获取的数据等。本地存储器119还可以包含用于存储微码的只读存储器(ROM)。虽然图1中的实例存储器子系统110已示为包含控制器115,但在本公开的另一实施例中,存储器子系统110可不包含控制器115,而是可以依赖于外部控制(例如,由外部主机或由与存储器子系统分开的处理器或控制器提供的外部控制)。

[0025] 通常,控制器115可以从主机系统120接收命令或操作,并且可以将命令或操作转换成指令或适当的命令,以实现对于存储器组件109A到109N的期望的访问。控制器115可以负责其它操作,例如耗损均衡操作、垃圾收集操作、错误检测和纠错码(ECC)操作、加密操作、高速缓存操作,以及与存储器组件109A到109N相关联的逻辑块地址与物理块地址之间的地址转换。控制器115可进一步包含主机接口电路系统以经由物理主机接口与主机系统120通信。主机接口电路系统可以将从主机系统接收到的命令转换成命令指令以访问存储器组件109A到109N,以及将与存储器组件109A到109N相关联的响应转换成用于主机系统120的信息。

[0026] 存储器子系统110还可以包含未示出的附加电路系统或组件。在一些实施例中,存储器子系统110可以包含高速缓存或缓冲(例如DRAM)和地址电路系统(例如行解码器和列解码器),其可从控制器115接收地址且对地址进行解码以访问存储器组件109A到109N。

[0027] 存储器子系统110包含缓冲区管理器113,其可以将请求缓冲区的一部分用于读请求以及写请求。在一些实施例中,存储器子系统110中的控制器115包含缓冲区管理器113的至少一部分。例如,控制器115可以包含实施缓冲区管理器113的逻辑电路系统。例如,控制器115的处理装置117(处理器)可经配置以执行存储在本地存储器中的指令,以执行本文描述的缓冲区管理器113的操作。在一些实施例中,缓冲区管理器113是主机系统120的操作系统的一部分、装置驱动器或应用程序。

[0028] 主机系统120包含请求管理器112,所述请求管理器可以跟踪存储器子系统110中的可用缓冲区容量,以向存储器子系统110传输读请求和写请求。在一些实施例中,主机系统120中的控制器116包含请求管理器112的至少一部分。例如,控制器116可以包含实施请求管理器112的逻辑电路系统。例如,主机系统120的处理装置118(处理器)可经配置以执行存储在本地存储器中的指令,以执行本文描述的请求管理器112的操作。在一些实施例中,请求管理器112是主机系统120的操作系统的一部分、装置驱动器或应用程序。

[0029] 存储器子系统110的缓冲区管理器113可以使用共享缓冲区来缓冲读请求和写请求。缓冲区共享用于读请求和写请求。可选地,缓冲区管理器113可以进一步使用只读缓冲区,所述只读缓冲区仅用于缓冲读请求而不用于缓冲写请求。类似地,缓冲区管理器113可

以可选地进一步使用只写缓冲区,所述只写缓冲区仅用于缓冲写请求而不适用于缓冲读请求。在某些情况下,缓冲区管理器113可以将缓冲区动态地划分为共享缓冲区、只读缓冲区和只写缓冲区。在完成执行一或多个缓冲的写请求之后,经执行写请求所使用的缓冲区容量变得可用于接受来自主机系统120的新的写请求。缓冲区管理器113可以跟踪变得可用于接受新的写请求的回收的缓冲区容量。当存储器子系统110向主机系统120传送用于进一步写请求的可用缓冲区容量的量时,允许可用缓冲区的量来写入主机系统120可以发送的请求而不会引起缓冲区溢出。缓冲区管理器113扣除该量以获得在缓冲区管理器113处跟踪的剩余可用缓冲区容量。

[0030] 主机系统120的请求管理器112可以跟踪存储器子系统110的可用读缓冲区容量以接受进一步的读请求,跟踪存储器子系统110的可用写缓冲区容量以接受进一步的写请求,和/或跟踪存储器子系统110的总可用缓冲区容量以接受进一步的读请求和写请求。当主机系统120向存储器子系统110发送写请求时,请求管理器112将可用写缓冲区容量和总可用缓冲区容量减少用于写请求的缓冲区容量的量。当主机系统120向存储器子系统110发送读请求时,请求管理器112将可用读缓冲区容量和总可用缓冲区容量减少用于读请求的缓冲区容量的量。当存储器子系统110向主机系统120发送读请求所请求的数据时,请求管理器112将可用读缓冲区容量和总可用缓冲区容量增加用于读请求的缓冲区容量的量。当存储器子系统110向主机系统120传送用于进一步写请求的可用缓冲区容量的量时,请求管理器112将可用写缓冲区容量和总可用缓冲区容量增加由存储器子系统110指定的量。因此,当可用写缓冲区容量小于写请求所需的缓冲区空间量时,请求管理器112推迟传输写请求,直到可用写缓冲区容量大于写请求所需的缓冲区空间量。类似地,当可用读缓冲区容量小于读请求所需的缓冲区空间量时,请求管理器112推迟传输读请求,直到可用读缓冲区容量大于读请求所需的缓冲区空间量。

[0031] 下面描述关于缓冲区管理器113和请求管理器112的操作的更多细节。

[0032] 图2示出根据本公开的一些实施例的包含请求管理器112和缓冲区管理器113的实例计算系统。

[0033] 描述图2时仅出于说明的非限制性目的,主机系统120的控制器116在下文有时被称为存储器控制器116,且存储器子系统110的控制器115在下文有时被称为媒体控制器115。

[0034] 在图2中,主机系统120与存储器子系统110之间的通信通道包含命令总线121、数据总线123、事务总线125和元数据总线127。用于通信通道的通信协议允许异步访问存储器子系统110以由主机系统120进行数据存储和检索。例如,存储器子系统110可以是NVDIMM;并且主机系统120可以使用命令总线121、数据总线123、事务总线125和元数据总线127根据JEDEC NVDIMM-P总线协议访问存储器控制器116。

[0035] 例如,存储器控制器116可以发出将数据存储存储在存储器子系统110中的写命令。从在命令总线121上传输写命令开始的固定和预定时间窗口之后,存储器控制器116开始在数据总线123上传输数据。不需要存储器子系统110在预定时间段内完成写命令的操作。这种写命令的实例包含在JEDEC NVDIMM-P总线协议中标识的XWRITE和PWRITE。

[0036] 例如,存储器控制器116可以发出请求来自存储器子系统110的信息的读命令。不需要存储器子系统110在从读命令开始的预定时间窗口内生成响应。这种读命令的实例包

含在JEDEC NVDIMM-P总线协议中标识的XREAD和SREAD。可以给定XREAD预定的读ID,以指示其是将返回系统状态但不会直接访问媒体的信息请求(status_read)。

[0037] 响应于读命令,存储器子系统110准备读命令所请求的数据。例如,媒体控制器115可以从媒体(例如109A、……、或109N)检索数据,并将检索数据缓冲在本地存储器119或另一存储器中,使得当请求此类传输时可以在预定时间窗口内将所述数据可以成功地传输到存储器控制器116。

[0038] 当所请求的数据准备好传输时,存储器子系统110可以在事务总线125中提供响应信号。当向存储器控制器116通知存储器子系统110准备传输某些数据时,存储器控制器116可以提供发送命令,以请求存储器子系统110在从发送命令开始的预定时间窗口内开始在数据总线123上传输数据。当对发送命令作出响应时,存储器子系统115还可以发送事务状态信息,例如标识对应读命令的读ID、如下文进一步论述的写积分信息、与事务相对应的元数据,和/或纠错码(ECC)。此类发送命令的实例是在JEDEC NVDIMM-P总线协议中标识的SEND。

[0039] 存储器子系统110可以将命令总线121接收的读命令和写命令缓冲在本地存储器119或另一存储器中。媒体控制器115可以按与接收命令的顺序不同的顺序执行缓冲的命令。

[0040] 存储器子系统110具有一定的容量,用于缓冲待定的读命令和写命令及其相关联的数据。存储器控制器116和媒体控制器115可以彼此通信以防止存储器子系统110中的缓冲区溢出。

[0041] 例如,写积分可以用于表示可用于缓冲写命令及其预定大小的关联数据的缓冲区容量单位。在一些情况下,写命令可以具有大于预定大小的数据;并且此类写命令需要多个写积分来将命令及其数据缓存在存储器子系统110中。

[0042] 存储器控制器116可以保持其可以用来在命令总线121上将写命令传输到存储器子系统110的写积分的计数。当通过命令总线121发送写命令时,存储器控制器116扣除写命令所使用的写积分。为了避免缓冲区溢出,当存储器控制器11没有足够的写积分可用于将写命令传输到存储器子系统110时,存储器控制器116不应传输写命令。

[0043] 媒体控制器115可以保持其可以返回到存储器控制器116以完成写命令的写积分的计数。在存储器子系统110中缓冲的写命令完成之后,写命令所使用的缓冲区空间可以被释放以接受来自存储器控制器116的进一步的写命令。已完成的写命令所使用的写积分可以被加到可被返回到存储器控制器116的写积分的计数中。

[0044] 存储器子系统110可以使用元数据总线127来指定其返回到存储器控制器116的写积分的数目。例如,在事务总线125上发送响应信号以使存储器控制器116能够发出发送命令之后,媒体控制器115可以使用元数据总线127传输返回的写积分的数目。存储器子系统110可以响应于读命令(例如JEDEC NVDIMM-P总线协议中标识的XREAD和SREAD)传输此类响应信号。响应信号的实例是在JEDEC NVDIMM-P总线协议中标识的RSPx_n。

[0045] 当存储器控制器116使用读命令来请求从地址中检索数据时,存储器控制器116可以在读命令之后紧接放置地址命令以指定地址。类似地,当存储器控制器116使用写命令将数据存储在地地址处时,存储器控制器116可以在写命令之后紧接放置地址命令以指定地址。此类地址命令的实例是在JEDEC NVDIMM-P总线协议中标识的XADR。

[0046] 存储器子系统110可以包含用于存储从主机系统120接收的待定的读写命令/请求的共享缓冲区。共享缓冲区可以在本地存储器119或存储器子系统110的另一存储器中实施。

[0047] 存储器子系统110可以包含用于存储从主机系统120接收的待定的读命令/请求的读缓冲区。读缓冲区可以在本地存储器119或存储器子系统110的另一存储器中实施。写命令/请求及其数据不存储在读缓冲区中。

[0048] 存储器子系统110可以包含用于存储从主机系统120接收的待定的写命令/请求的写缓冲区。写缓冲区可以在本地存储器119或存储器子系统110的另一存储器中实施。读命令/请求及其数据不存储在写缓冲区中。

[0049] 缓冲区管理器113可以管理共享缓冲区、读缓冲区和/或写缓冲区的使用。缓冲区管理器113跟踪表示从完成的写请求释放的缓冲区容量的回收的写积分。缓冲区管理器113可以将回收的写积分作为写积分增量返回给主机系统120。请求管理器112可以在主机系统120可用的读缓冲区和/或共享缓冲区中跟踪表示读缓冲区容量的读积分,以向存储器子系统110发送在主机系统120指定的地址处检索数据的读请求。请求管理器112可以在主机系统120可用的写缓冲区和/或共享缓冲区中跟踪表示写缓冲区容量的写积分,以向存储器子系统110发送在主机系统120指定的地址处存储数据的写请求。请求管理器112可以在主机系统120可用的读缓冲区、写缓冲区和/或共享缓冲区中跟踪表示总缓冲区容量的总积分以发送读请求和/或写请求。

[0050] 图3A-3C示出根据本公开的一些实施例的实例缓冲区配置。

[0051] 图3A-3B中的专用读缓冲区201有多个读积分。每个读积分表示用于缓冲读请求及其从读请求的地址检索的数据的预定量的读缓冲区容量。写请求及其数据不可存储在读缓冲区201中。

[0052] 图3A-3B中的专用写缓冲区203有多个写积分。每个写积分表示用于缓冲写命令及其要存储在写请求的地址中的数据的数据的预定量的写缓冲区容量。读请求及其数据不可存储在写缓冲区203中。

[0053] 图3B-3C中的共享缓冲区205有多个共享积分。共享缓冲区205可以用于存储写请求及其数据、读请求及其数据,或读请求和写请求及其数据的组合。表示预定量的写缓冲区容量的每个共享积分可以用作读积分或写积分。

[0054] 图3A示出其中存储器子系统110包含专用读缓冲区201和专用写缓冲区203的缓冲区配置。总积分T等于W个写积分和R个读积分之和。

[0055] 图3B示出其中存储器子系统110包含专用读缓冲区201、专用写缓冲区203和共享缓冲区205的缓冲区配置。读积分R的数目等于专用读缓冲区201和共享缓冲区205的组合容量;并且写积分W的数目等于专用写缓冲区203和共享缓冲区205的组合容量。总积分T是专用读缓冲区201、专用写缓冲区203和共享缓冲区205的组合容量。总积分T等于W个写积分和R个读积分之和减去重叠的S个共享积分。

[0056] 基于已知的共享缓冲区205的容量S,可以根据可用总积分t和可用写积分w计算用于传输进一步的读请求的可用读积分r。例如, $r=t-\min(w,S)$ 。

[0057] 基于已知的共享缓冲区205的容量S,可以根据可用总积分t和可用读积分r计算用于传输进一步的写请求的可用写积分w。例如, $w=t-\min(r,S)$ 。

[0058] 基于已知的共享缓冲区205的容量S,可以根据可用写积分w和可用读积分r计算用于传输进一步的写/读请求的可用总积分t。例如, $t = w + r - \min(r, w, S)$ 。

[0059] 图3C示出其中存储器子系统110包含共享缓冲区205但是不包含专用读缓冲区或写缓冲区的缓冲区配置。在图3C的实例中,读积分的数目R、写积分的数目W和积分的总数目T等于共享缓冲区205的容量S。用于传输进一步的写/读请求的可用总积分t、可用写积分w、可用读积分r彼此相等。

[0060] 当存储器子系统110从主机系统120接收到读请求时,缓冲区管理器113可以在读缓冲区201可用时将读请求存储在所述读缓冲区中,或者可以在读缓冲区201中没有可用的容量时将读请求存储在共享缓冲区205中。在媒体控制器115检索到读请求所请求的数据并缓冲所检索的数据之后,媒体控制器115可以发送响应信号以使存储器控制器116从媒体控制器115接收数据。在检索数据已从存储器子系统110传输到主机系统120之后,可以从读缓冲区201或共享缓冲区205清除读请求及其数据。

[0061] 当存储器子系统110从主机系统120接收到写请求时,缓冲区管理器113可以在写缓冲区203可用时将写请求及其数据存储在所述写缓冲区中,或者可以在写缓冲区203中没有可用的容量时将写请求存储在共享缓冲区205中。在媒体控制器115完成将写请求的数据存储在存储器组件109A到109N中之后,可以从写缓冲区203或共享缓冲区205清除写请求及其数据。

[0062] 在一些实施例中,可以省略共享缓冲区205(例如,如图3A所示)。在其它实施例中,可以省略专用读缓冲区201和/或专用写缓冲区203(例如,如图3B和3C所示)。专用读缓冲区201和/或专用写缓冲区203可以具有不同的容量。缓冲区管理器113可以将缓冲存储器动态地划分为读缓冲区201、写缓冲区203和/或共享缓冲区205。

[0063] 缓冲区管理器113可以对从存储器子系统110中完成的写请求中回收的写积分进行计数,并将回收的写积分分配为写积分增量,从存储器子系统110传输到主机系统120。例如,根据JEDEC NVDIMM-P总线协议,响应于来自主机系统120的发送命令,可以使用元数据总线127来传输写积分增量。

[0064] 请求管理器116可以跟踪用于传输进一步的读请求的可用读积分、用于传输进一步的写请求的可用写积分和/或用于传输进一步的读/写请求的可用总积分。

[0065] 当读请求被传输到存储器子系统110时,请求管理器116可以递减可用读积分和/或可用总积分。当从存储器子系统110接收到读请求中所请求的数据时,请求管理器116可以递增可用读积分和/或可用总积分。

[0066] 当写请求被传输到存储器子系统110时,请求管理器116可以递减可用写积分和/或可用总积分。根据从存储器子系统110传输到主机系统120的写积分增量,请求管理器116可以递增可用写积分和/或可用总积分。

[0067] 此外,当可用读积分减少至共享缓冲区205的总容量S时,对于读请求的进一步减少将使得可用写积分相等地减少。在这种情况下,将共享缓冲区205用于读请求也减少了共享缓冲区205用于接收写请求的可用容量。

[0068] 类似地,当可用写积分减少至共享缓冲区205的总容量S时,对于写请求的进一步减少将使得可用读积分相等地减少。在这种情况下,将共享缓冲区205用于写请求也减少了共享缓冲区205用于接收写请求的可用容量。

[0069] 替代地,请求管理器116不减少共享缓冲区205中读请求所使用的写积分,也不减少共享缓冲区205中写请求所使用的读积分。请求管理器116跟踪可用读积分、可用写积分和可用总积分。为了防止缓冲区溢出,当读请求的大小超过可用读积分或可用总积分时,请求管理器116不传输读请求。类似地,当写请求的大小超过可用写积分或可用总积分时,请求管理器116不传输写请求。

[0070] 例如,请求管理器116可以在不考虑写请求所使用的共享缓冲区205的情况下跟踪可用读积分 r ;并且请求管理器116可以在不考虑读请求所使用的共享缓冲区205的情况下跟踪可用写积分 w 。因此,写请求所使用的共享缓冲区205可以计算为 $\min(S-w, 0)$;并且读请求所使用的共享缓冲区205可以计算为 $\min(S-r, 0)$ 。因此,可用读积分的实际量为 $r - \min(S-w, 0)$;并且可用写积分的实际量为 $w - \min(S-r, 0)$ 。

[0071] 请求管理器116可经配置以取决于缓冲区配置来跟踪特定类型的积分并进行计数。例如,在存储器子系统110包含专用读缓冲区201和专用写缓冲区203但不包含共享缓冲区的情况下,如图3A所示,请求管理器116可以跟踪可用读积分和可用写积分,而无需跟踪可用总积分。替代地,请求管理器116可以跟踪可用总积分和可用写积分,并根据可用总积分与可用写积分之间的差来计算可用读积分。

[0072] 在另一实例中,当存储器子系统110包含共享缓冲区205但是不包含专用读缓冲区或专用写缓冲区时(或者当读缓冲区201和写缓冲区203已满时),如图3C所示,请求管理器116可以跟踪可用读积分、可用写积分或可用总积分中的单个积分,因为它们彼此相等。

[0073] 图4-9示出的序列图示出根据本公开的一些实施例的结合读操作或写操作的缓冲区容量跟踪的实例。缓冲区管理器113和请求管理器112可以使用图4-9所示的序列的组合。

[0074] 参考图1和2所示的主机系统120与存储器子系统110之间的通信以及由所述主机系统和所述存储器子系统进行的操作结合图3A-3C所示的缓冲区配置描述图4-9。主机系统120的一些或全部操作一般可以由存储器控制器116执行,并且具体地说至少部分地由请求管理器112执行。存储器子系统110的一些或全部操作可以一般可以由媒体控制器115执行,并且具体地说可以至少部分地由缓冲区管理器113执行。

[0075] 图4示出可用读积分的跟踪。

[0076] 在操作410中,存储器子系统110传输并且主机系统120接收存储器子系统110中可用以接受来自主机系统120的读请求的读积分的初始数量的指示。替代地,主机系统120可以从除存储器子系统110以外的其它源(例如从在主机系统120上存储并运行的装置驱动器、存储用于存储器子系统110的配置参数的存储器位置等)接收读积分的初始数量的指示。

[0077] 在操作412中,主机系统120将可用读积分的数量设置为等于在操作410中接收到的初始数量。结果,主机系统120(例如,其请求管理器112)可以知道存储器子系统110在缓冲读请求及其数据时具有的缓冲区容量的总量。

[0078] 在操作414中,主机系统120传输并且存储器子系统110接收对存储在存储器子系统110的媒体或存储器组件109A到109N中的指定数据的读请求。读请求的实例是在JEDEC NVDIMM-P总线协议中标识的SREAD或XREAD命令以及相关联的XADR命令。读请求可以包含在所请求的数据被传输到主机系统120之前在存储器子系统110中可能需要的所请求的数据的大小指示。

[0079] 当在主机系统120处跟踪的当前可用读积分不小于所请求的数据的大小时,主机系统120传输读请求。如在主机系统120处跟踪的当前可用读积分小于所请求的数据的大小,则可以阻止主机系统120在操作414中传输读请求。

[0080] 在操作416中,主机系统120响应于在操作414中传输读请求而递减其可用读积分的数量。递减的数量是根据所请求的数据的大小。

[0081] 在操作418中,存储器子系统110缓冲读请求(例如,在读缓冲区201和/或共享缓冲区205中)。

[0082] 在操作420中,存储器子系统110从存储器子系统110的媒体检索所请求的数据。

[0083] 在操作422中,存储器子系统110缓冲所请求的数据以传输到主机系统120(例如,在读缓冲区201和/或共享缓冲区205中)。

[0084] 在操作424中,存储器子系统110传输并且主机系统120接收响应信号,例如在JEDEC NVDIMM-P总线协议中标识的RSPx_n信号。

[0085] 在操作426中,主机系统120传输并且存储器子系统110接收发送命令,所述发送命令请求存储器子系统110传输由响应信号指示为准备好传输的数据。发送命令的实例是在JEDEC NVDIMM-P总线协议中标识的SEND命令。

[0086] 在操作428中,存储器子系统110传输并且主机系统120接收在操作414的读请求中所请求的数据。作为传输所请求的数据的一部分或传输所请求的数据之后,存储器子系统110可以清除缓冲区(例如201和/或205)中的读请求和在读请求中所请求的数据。

[0087] 在操作432中,主机系统120响应于在操作426中接收到所请求的数据而递增可用读积分的数量。所请求的数据的接收向主机系统120指示用于读请求的缓冲区容量已被释放并且可获得/可用于进一步的读请求。

[0088] 通常,在操作414的读请求的传输之前,在操作414的读请求与操作420的所请求的数据的检索之间,在操作420的所请求的数据的检索与操作424的响应信号之间,以及在操作424的响应信号与操作426的发送命令之间等,主机系统120和存储器子系统110可以具有其它通信。此类通信可以包含另一读请求、写请求、状态查询等。

[0089] 可以针对每个读请求重复图4所示的某些操作,使得主机系统120在读请求被传输和服务时连续跟踪可用读积分。例如,可以针对每个读请求执行操作414到432,并且如上所述,主机系统120仅在有足够的读积分可用于读请求时才可以传输读请求。

[0090] 在图4中,主机系统120显式地跟踪读积分。这样,主机系统120仅在有足够的读积分可用于读请求时才发送读请求。另外,存储器子系统110不必显式地跟踪读积分。在其它实施例中,存储器子系统110可以可选地跟踪读积分并且可以将有关可用读积分的数量的信息提供给主机系统120。此外,在实施例中,存储器子系统110可经配置以在当前没有足够的读缓冲区容量可用于读请求的情况下拒绝读请求;以及在有足够的读缓冲区容量可用于读请求的情况下接受读请求。

[0091] 图5示出对可用写积分的跟踪。

[0092] 在操作510中,存储器子系统110传输并且主机系统120接收存储器子系统110中可用以接受来自主机系统120的写请求的写积分的初始数量的指示。替代地,主机系统120可以从除存储器子系统110以外的其它源(例如从在主机系统120上存储并运行的装置驱动器、存储用于存储器子系统110的配置参数的存储器位置等)接收写积分的初始数量的指

示。

[0093] 在操作512中,主机系统120将可用写积分的数量设置为等于在操作510中接收到的初始数量。结果,主机系统120(例如,其请求管理器112)可以知道存储器子系统110在缓冲写请求及其数据时具有的缓冲区容量的总量。

[0094] 在操作514中,主机系统120传输并且存储器子系统110接收对将指定数据存储到存储器子系统110的媒体或存储器组件109A到109N的写请求。写请求的实例是在JEDEC NVDIMM-P总线协议中标识的WRITE、XWRITE、PWRITE命令和相关联的XADR命令。写请求要存储的数据确定写请求的大小。

[0095] 当在主机系统120处跟踪的当前可用写积分不小于写请求的大小时,主机系统120传输写请求。如在主机系统120处跟踪的当前可用写积分小于写请求的大小,则可以阻止主机系统120在操作514中传输写请求。

[0096] 在操作516中,主机系统120响应于在操作514中传输写请求而递减其可用写积分的数量。递减的数量是根据写请求的大小。

[0097] 在操作518中,存储器子系统110缓冲写请求(例如,在读缓冲区201和/或共享缓冲区205中)。

[0098] 在操作520中,存储器子系统110将数据写入到存储器子系统110的媒体。

[0099] 在操作522,存储器子系统110清除写请求和从缓冲区(例如,在读缓冲区201和/或共享缓冲区205中)写入的数据。

[0100] 在操作524中,存储器子系统110响应于完成在操作514中接收到的写请求而递增其可用写积分的数量。递增的数量是根据完成的写请求的大小。

[0101] 在操作526中,主机系统120传输并且存储器子系统110接收请求存储器子系统110提供包含写积分的信息的读命令。读命令的实例是在JEDEC NVDIMM-P总线协议中标识的SREAD和XREAD(和status_read)。

[0102] 在操作424中,存储器子系统110传输并且主机系统120接收响应信号,例如在JEDEC NVDIMM-P总线协议中标识的RSPx_n信号。

[0103] 在操作426中,主机系统120传输并且存储器子系统110接收发送命令,所述发送命令请求存储器子系统110传输由响应信号指示为准备好传输的数据。发送命令的实例是在JEDEC NVDIMM-P总线协议中标识的SEND命令。

[0104] 在操作528中,存储器子系统110传输并且主机系统120接收写积分。当响应信号是针对使存储器子系统110从存储器子系统110的媒体中检索数据的读请求时,存储器子系统110还传输并且主机系统120接收在操作414的读请求中所请求的数据,如图4所示的操作428。在操作528中从存储器子系统110发送到主机系统的写积分可以包含作为完成操作514的写请求和/或另一写请求的结果的写积分增量。

[0105] 在操作530中,存储器子系统110根据在操作528中传输到主机系统120的写积分来递减其可用写积分的数量。

[0106] 在操作532中,主机系统120根据在操作528中传输到主机系统120的写积分来递增其可用写积分的数量。

[0107] 通常,在操作514的写请求的传输之前,在操作514的写请求与操作520的数据的写入之间,在操作520的数据的写入与操作424的响应信号之间,以及在操作424的响应信号与

操作426的发送命令之间等,主机系统120和存储器子系统110可以具有其它通信。此类通信可以包含另一写请求、读请求、状态查询等。

[0108] 可以针对每个写请求重复图5所示的某些操作,使得主机系统120在写请求被传输和服务时且在写积分被传输回主机系统时连续跟踪可用写积分。例如,可以针对每个写请求执行操作514到532,并且如上所述,主机系统120仅在有足够的写积分可用于写请求时才可以传输写请求。

[0109] 在图5中,存储器子系统110显式地跟踪根据完成的写请求生成的写积分增量,并且响应于来自主机系统120的发送命令而将更新写积分增量发送到主机系统120。当没有足够的写积分可用于写请求时,存储器子系统110可以拒绝写请求。在一些实施例中,存储器子系统110可以向主机系统120标识完成的写请求,使得主机系统120可以针对完成的写请求递增其写积分。

[0110] 图6示出结合读请求的可用总积分的跟踪。

[0111] 在操作610中,存储器子系统110传输并且主机系统120接收存储器子系统110中可用以接受来自主机系统120的读/写请求的总积分的初始数量的指示。替代地,主机系统120可以从除存储器子系统110以外的其它源(例如从在主机系统120上存储并运行的装置驱动器、存储用于存储器子系统110的配置参数的存储器位置等)接收总积分的初始数量的指示。

[0112] 在操作612中,主机系统120将可用总积分的数量设置为等于在操作610中接收到的初始数量。结果,主机系统120(例如,其请求管理器112)可以知道存储器子系统110在缓冲读/写请求及其数据时具有的缓冲区容量的总量。

[0113] 在操作414中,主机系统120传输并且存储器子系统110接收对存储在存储器子系统110的媒体或存储器组件109A到109N中的指定数据的读请求,如图4所示。

[0114] 在操作616中,主机系统120响应于在操作414中传输读请求而递减其可用总积分的数量。递减的数量是根据所请求的数据的大小。

[0115] 在操作620中,存储器子系统110可选地为读请求分配共享缓冲区空间(例如,当读缓冲区201已满或未在存储器子系统110中配置时)。

[0116] 在操作418到428中,存储器子系统110完成读请求并将所请求的数据提供给主机系统120,如图4所示。

[0117] 在操作632中,主机系统120响应于在操作426中接收到所请求的数据而递增可用读积分的数量。所请求的数据的接收向主机系统120指示用于读请求的缓冲区容量已经被释放并且可获得/可用于进一步的读/写请求。

[0118] 通常,在操作414的读请求之前,在操作414的读请求与操作420的所请求的数据的检索之间,在操作420的所请求的数据的检索与操作424的响应信号之间,以及在操作424的响应信号与操作426的发送命令之间等,主机系统120和存储器子系统110可以具有其它通信。此类通信可以包含另一读请求、写请求、状态查询等。

[0119] 在一些情况下,总积分的至少一部分可用于读请求或写请求。例如,总积分的这一部分可以来自共享缓冲区205。例如,当读缓冲区201和写缓冲区203均未满时,总积分的这一部分可以来自读缓冲区201或来自写缓冲区203。在这种情况下,当其总积分数量不小于读请求的大小时,主机系统120可以在操作414中传输读请求。当在主机系统120处跟踪的总

积分的数量小于读请求的大小时,主机系统120不传输读请求。

[0120] 图7示出结合写请求的可用总积分的跟踪。

[0121] 在操作610和612中,主机系统120设置可用总积分的初始数量,如图6所示。

[0122] 在操作514中,主机系统120传输并且存储器子系统110接收对将指定数据存储到存储器子系统110的媒体或存储器组件109A到109N的写请求,如图5所示。

[0123] 在操作716中,主机系统120响应于在操作514中传输写请求而递减其可用总积分的数量。递减的数量是根据写请求的大小。

[0124] 在操作720中,存储器子系统110可选地为写请求分配共享缓冲区空间(例如,当读缓冲区201已满或未在存储器子系统110中配置时)。

[0125] 在操作518到530中,存储器子系统110处理写请求,并将写积分发送回主机系统120,如图5所示。

[0126] 在操作732中,主机系统120响应于在操作528中接收到写积分而递增可用读积分的数量。

[0127] 通常,在操作514的写请求之前,在操作514的写请求与操作520的数据的写入之间,在操作520的数据的写入与操作424的响应信号之间,以及在操作424的响应信号与操作426的发送命令之间等,主机系统120和存储器子系统110可以具有其它通信。此类通信可以包含另一写请求、读请求、状态查询等。

[0128] 在一些情况下,总积分的至少一部分可用于读请求或写请求。例如,总积分的这一部分可以来自共享缓冲区205。例如,当读缓冲区201和写缓冲区203均未满时,总积分的这一部分可以来自读缓冲区201或来自写缓冲区203。在这种情况下,当其总积分数量不小于写请求的大小时,主机系统120可以在操作514中传输写请求。当在主机系统120处跟踪的总积分的数量小于写请求的大小时,主机系统120不传输写请求。

[0129] 在一些情况下,如图4中所跟踪的可用读积分或如图6和7所跟踪的可用总积分可能不足以确定主机系统120是否可以发送读请求。主机系统120可以跟踪可用读积分和可用总积分两者,如下文进一步论述。

[0130] 图8示出结合读请求的可用总积分和可用读积分的跟踪。

[0131] 在操作810中,存储器子系统110传输并且主机系统120接收存储器子系统110中可用以接受来自主机系统120的读/写请求的总积分的初始数量的指示,以及存储器子系统110中可用以接受来自主机系统120的读请求的读积分的初始数量的指示。替代地,主机系统120可以从除存储器子系统110以外的其它源(例如从在主机系统120上存储并运行的装置驱动器、存储用于存储器子系统110的配置参数的存储器位置等)接收总积分和读积分的初始数量的指示。

[0132] 在操作412和612中,主机系统120将可用读积分的数量设置为等于在操作810中接收的初始数量,并且将可用总积分的数量设置为等于在操作810中接收的初始数量。

[0133] 在操作414中,主机系统120传输并且存储器子系统110接收对存储在存储器子系统110的媒体或存储器组件109A到109N中的指定数据的读请求,如图4所示。

[0134] 在操作416和616中,主机系统120递减读积分和总积分,如图4和6所示。

[0135] 在操作812中,存储器子系统110确定读缓冲区201是否有用于读请求的足够容量;如果没有,则在操作820中,存储器子系统110为读请求分配共享缓冲区容量。

[0136] 在操作418到428中,存储器子系统110完成读请求并将所请求的数据提供给主机系统120,如图4和6所示。

[0137] 在操作432和632中,主机系统120响应于在操作426中接收到所请求的数据而递增可用读积分的数量和可用总积分的数量,如图4和6所示。

[0138] 在图8中,当其可用总积分的数量和其可用读积分的数量均不小于读请求的大小时,主机系统120可以在操作414中传输读请求。当在主机系统120处跟踪的总积分的数量小于读请求的大小,或者在主机系统120处跟踪的读积分的数量小于读请求的大小时,主机系统120不传输读请求。

[0139] 图9示出结合写请求的可用总积分和可用读积分的跟踪。

[0140] 在操作910中,存储器子系统110传输并且主机系统120接收存储器子系统110中可用以接受来自主机系统120的读/写请求的总积分的初始数量的指示,以及存储器子系统110中可用以接受来自主机系统120的写请求的写积分的初始数量的指示。替代地,主机系统120可以从除存储器子系统110以外的其它源(例如从在主机系统120上存储并运行的装置驱动器、存储用于存储器子系统110的配置参数的存储器位置等)接收总积分和读积分的初始数量的指示。

[0141] 在操作512和612中,主机系统120将可用写积分的数量设置为等于在操作910中接收的初始数量,并且将可用总积分的数量设置为等于在操作910中接收的初始数量。

[0142] 在操作514中,主机系统120传输并且存储器子系统110接收对将指定数据存储到存储器子系统110的媒体或存储器组件109A到109N的写请求,如图5所示。

[0143] 在操作516和716中,主机系统120递减读积分和总积分,如图5和7所示。

[0144] 在操作912中,存储器子系统110确定写缓冲区203是否有用于写请求的足够容量;如果没有,则在操作920中,存储器子系统110为写请求分配共享缓冲区容量。

[0145] 在操作518到530中,存储器子系统110处理写请求,并将写积分发送回主机系统120,如图5和7所示。

[0146] 在操作532和732中,主机系统120响应于在操作528中接收到写积分而递增可用写积分和可用总积分的数量。

[0147] 在图9中,当其可用总积分的数量和其可用写积分的数量均不小于写请求的大小时,主机系统120可以在操作514中传输读请求。当在主机系统120处跟踪的总积分的数量小于写请求的大小,或者在主机系统120处跟踪的写积分的数量小于写请求的大小时,主机系统120不传输写请求。

[0148] 在一些情况下,当主机系统120具有初始总积分数量 T 、读积分数量 R 和写积分数量 W 的指示时,主机系统120可以根据 $S=W+R-T$ 来计算共享缓冲区205的容量 S 。替代地,可以根据存储在计算机系统配置参数中的配置参数来确定共享缓冲区205的容量 S 。

[0149] 随后,主机系统120的请求管理器112可以基于将可用读积分 r 与共享缓冲区205的容量 S 进行比较来确定读缓冲区201是否已满。如果 $r < S$,则读缓冲区201已满;并且用于读请求的附加容量是来自共享缓冲区205,这减少了可用写积分。当可用写积分的数量由于读请求对共享缓冲区205的使用而减少时,可用写积分的数量反映了可用于传输进一步写请求的实际写缓冲区容量;并且不必像图9中那样单独跟踪总积分。

[0150] 随后,主机系统120的请求管理器112可以基于将可用写积分 w 与共享缓冲区205的

容量S进行比较来确定写缓冲区203是否已满。如果 $w < S$,则写缓冲区203已满;并且用于写请求的附加容量是来自共享缓冲区205,这减少了可用读积分。当可用读积分的数量由于写请求对共享缓冲区205的使用而减少时,可用读积分的数量反映了可用于传输进一步读请求的实际读缓冲区容量;并且不必像图8中那样单独跟踪总积分。

[0151] 当主机系统120具有读积分增量时(例如,响应于接收到读请求所请求的数据),主机系统120的请求管理器112可以确定增加的容量是否在共享缓冲区205中。在当前读积分量小于S时,可用读积分增加到S是处于共享缓冲区205。共享缓冲区205中增加的读积分也增加了可用写积分。

[0152] 类似地,当主机系统120具有写积分增量时(例如,响应于从存储器子系统110接收到写积分),主机系统120的请求管理器112可以确定增加的容量是否在共享缓冲区205中。在当前写积分量小于S时,可用写积分增加到S是处于共享缓冲区205。共享缓冲区205中增加的写积分也增加了可用读积分。

[0153] 图10是根据本公开的一些实施例的用于管理请求请求和写请求共享的缓冲区的实例方法的流程图。图10的方法可以由处理逻辑来执行,所述处理逻辑可以包含硬件(例如,处理装置、电路系统、专用逻辑、可编程逻辑、微码、装置的硬件、集成电路等)、软件(例如,在处理装置上运行或执行的指令),或其组合。在一些实施例中,图10的方法可以至少部分地由图1或2的缓冲区管理器113执行。尽管以特定的顺序或顺序示出,但除非另有说明,否则可以修改过程的顺序。因此,示出的实施例应理解为仅是举例说明,并且示出的过程可以以不同的顺序执行,并且一些过程可以并行地执行。另外,在各种实施例中可以省略一或多个过程。因此,并非在每个实施例中都需要所有过程。其它处理流程也是可能的。

[0154] 在框310,存储器子系统110的媒体控制器115从主机系统120接收写请求。

[0155] 在框312,在写请求待定期期间,缓冲区管理器113将写请求的数据存储在缓冲区的一部分(例如203和205)中。

[0156] 在框314,存储器子系统110的媒体控制器115从主机系统120接收读请求。

[0157] 在框316,缓冲区管理器113将读请求的数据存储在缓冲区的一部分(例如201和205)中,所述一部分与缓冲区的在写请求待定期期间存储写请求的数据的部分(例如203和205)至少部分地重叠。

[0158] 因此,缓冲区管理器113可以使用缓冲区的至少同一部分(例如205)来在缓冲区的不同操作时间段(例如205)缓冲读请求和写请求。

[0159] 图11和12是根据本公开的一些实施例的用于管理对存储器子系统的请求的实例方法的流程图。图11和12的方法可以由处理逻辑来执行,所述处理逻辑可以包含硬件(例如,处理装置、电路系统、专用逻辑、可编程逻辑、微码、装置的硬件、集成电路等)、软件(例如,在处理装置上运行或执行的指令),或其组合。在一些实施例中,图11或12的方法可以至少部分地由图1或2的请求管理器112执行。尽管以特定的顺序或顺序示出,但除非另有说明,否则可以修改过程的顺序。因此,示出的实施例应理解为仅是举例说明,并且示出的过程可以以不同的顺序执行,并且一些过程可以并行地执行。另外,在各种实施例中可以省略一或多个过程。因此,并非在每个实施例中都需要所有过程。其它处理流程也是可能的。

[0160] 在框330,请求管理器112存储存储器子系统110的可用缓冲区容量的总量,用于将读请求和写请求传输到存储器子系统110。

- [0161] 在框332,存储器控制器116将读请求传输到存储器子系统110。
- [0162] 在框334,请求管理器112基于读请求而减少总量。
- [0163] 在框336,存储器控制器116将写请求传输到存储器子系统110。
- [0164] 在框338,请求管理器112基于写请求而减少总量。
- [0165] 请求管理器112可以响应于接收到读请求所请求的数据而增加可用缓冲区容量的总量,并且根据从存储器子系统110传送的写积分增量来增加总量。请求管理器112可以基于当前总量是否不小于读请求或写请求的大小而确定是否传输读请求或写请求。
- [0166] 在框360,请求管理器112存储存储器子系统110的读缓冲区容量的量,用于将读请求传输到存储器子系统110。
- [0167] 在框362,存储器控制器116将读请求传输到存储器子系统110。
- [0168] 在框364,请求管理器112基于读请求而减少量。
- [0169] 在框366,存储器控制器116从存储器子系统110接收读请求所请求的数据。
- [0170] 在框368,请求管理器112基于接收到读请求所请求的数据而增加量。
- [0171] 请求管理器112可以基于可用读缓冲区容量的当前量是否不小于读请求的大小来确定是否传输读请求。
- [0172] 图13示出计算机系统200的实例机器,在所述实例机器内可执行用于使机器执行本文中所述的方法中的任何一或多种的指令集。在一些实施例中,计算机系统200可以对应于主机系统(例如,图1的主机系统120),所述主机系统包含、耦合到或利用存储器子系统(例如,图1的存储器子系统110)或可用于执行请求/缓冲区管理器213的操作(例如,执行指令以执行与参考图1-12描述的请求管理器112和/或缓冲区管理器113相对应的操作)。在替代实施例中,机器可以连接(例如联网)到LAN、内联网、外联网和/或互联网中的其它机器。机器可作为对等(或分散式)网络环境中的对等机器或作为云计算基础设施或环境中的服务器或客户端机器而以客户端-服务器网络环境中的服务器或客户端机器操作。
- [0173] 机器可以是个人计算机(PC)、平板PC、机顶盒(STB)、个人数字助理(PDA)、蜂窝电话、网络设备、服务器、网络路由器、开关或网桥、或能够(依序或以其它方式)执行一组指令的任何机器,所述指令指定要由所述机器采取的动作。另外,尽管示出了单个机器,但是术语“机器”也应被理解为包含机器的任何集合,所述机器单独地或共同地执行一(或多个)指令集以进行本文中所述的方法中的任何一或多种。
- [0174] 实例计算机系统200包含处理装置202、主存储器204(例如,只读存储器(ROM)、快闪存储器、动态随机存取存储器(DRAM)例如同步DRAM(SDRAM)或Rambus DRAM(RDRAM)等)、静态随机存取存储器(SRAM)等),以及数据存储系统218,其经由总线230(其可包含多个总线)彼此通信。
- [0175] 处理装置202表示一或多个通用处理装置,例如微处理器、中央处理单元(CPU)等。更具体地,处理装置可以是复杂指令集计算(CISC)微处理器、精简指令集计算(RISC)微处理器、超长指令字(VLIW)微处理器或实施其它指令集的处理器,或实施指令集的组分的处理器。处理装置202也可以是一或多个专用处理装置,例如专用集成电路(ASIC)、现场可编程门阵列(FPGA)、数字信号处理器(DSP)、图形处理单元(GPU)、网络处理器等。处理装置202经配置以执行指令226,用于执行本文讨论的操作和步骤。计算机系统200可以进一步包含网络接口装置208,以通过网络220进行通信。

[0176] 数据存储系统218可以包含机器可读存储媒体224(也称为计算机可读媒体),其上存储有一或多个指令集226或体现本文中所描述的方法或功能中的任何一或多种的软件。在计算机系统200执行指令期间,指令226还可以全部或至少部分地驻存在主存储器204内和/或处理装置202内,主存储器204和处理装置202也构成机器可读存储媒体。机器可读存储媒体224、数据存储系统218和/或主存储器204可以对应于图1的存储器子系统110。

[0177] 在一个实施例中,指令226包含用于实施与请求/缓冲区管理器213(例如,参考图1到12描述的请求管理器112和/或缓冲区管理器113)相对应的功能的指令。尽管在实例实施例中将机器可读存储媒体224展示为单个媒体,但术语“机器可读存储媒体”应被视为包含存储一或多个指令集的单个媒体或多个媒体。术语“机器可读存储媒体”还应被认为包含能够存储或编码供机器执行的指令集合且致使机器执行本公开的方法中的任何一种或多种的任何媒体。因此,应认为术语“机器可读存储媒体”包含但不限于固态存储器、光学媒体以及磁性媒体。

[0178] 已关于计算机存储器内的数据位的操作的算法和符号表示而呈现先前详细描述的一些部分。这些算法描述和表示是数据处理领域的技术人员用以将其工作的主旨最有效地传达给本领域的其它技术人员的方式。算法在这里并且通常被认为是产生期望的结果的操作的自洽序列。操作是要求对物理量进行物理控制的操作。这些量通常但未必呈能够被存储、组合、比较和以其它方式操控的电或磁信号的形式。有时,主要出于通用的原因,已经证明将这些信号称为位、值、元件、符号、字符、术语、数目或类似物是方便的。

[0179] 然而,应牢记,所有这些和类似术语应与适当物理量相关联,且仅仅是应用于这些量的方便标签。本公开可以指操控和变换计算机系统的寄存器和存储器内的表示为物理(电子)数量的数据为计算机系统存储器或寄存器或其它这类信息存储系统内的类似地表示为物理量的其它数据的计算机系统或类似电子计算装置的动作和过程。

[0180] 本公开还涉及用于执行本文中的操作的设备。此设备可以出于所需目的而专门构造,或其可以包含通过存储在计算机中的计算机程序选择性地激活或重新配置的通用计算机。此类计算机程序可存储在计算机可读存储媒体中,如但不限于任何类型的盘(包含软盘、光盘、CD-ROM和磁性光盘)、只读存储器(ROM)、随机存取存储器(RAM)、EPROM、EEPROM、磁卡或光卡或适合于存储电子指令的任何类型的媒体,它们各自耦合到计算机系统总线。

[0181] 本文中呈现的算法和显示器在本质上并不与任何特定计算机或其它设备相关。各种通用系统可与根据本文中的教示的程序一起使用,或其可证明构造用以执行所述方法更加专用的设备是方便的。将如下文描述中所阐述的那样来呈现多种这些系统的结构。此外,并不参考任何特定编程语言来描述本公开。应了解,可以使用各种编程语言来实施如本文中所描述的本公开的教示。

[0182] 本公开可提供为计算机程序产品或软件,其可包含在其上存储有可用于编程计算机系统(或其它电子装置)以进行根据本公开的过程的指令的机器可读媒体。机器可读媒体包含用于以机器(例如,计算机)可读的形式存储信息的任何机制。在一些实施例中,机器可读(例如计算机可读)媒体包含机器(例如计算机)可读存储媒体,例如只读存储器(“ROM”)、随机存取存储器(“RAM”)、磁盘存储媒体、光学存储媒体、闪存存储器组件等。

[0183] 在前述说明书中,已参考其特定实例实施例描述了本公开的实施例。应显而易见的是,可在不脱离如所附权利要求书中阐述的本公开的实施例的更广精神和范围的情况下

对本公开进行各种修改。因此,应在说明性意义上而非限制性意义上看待说明书和图式。

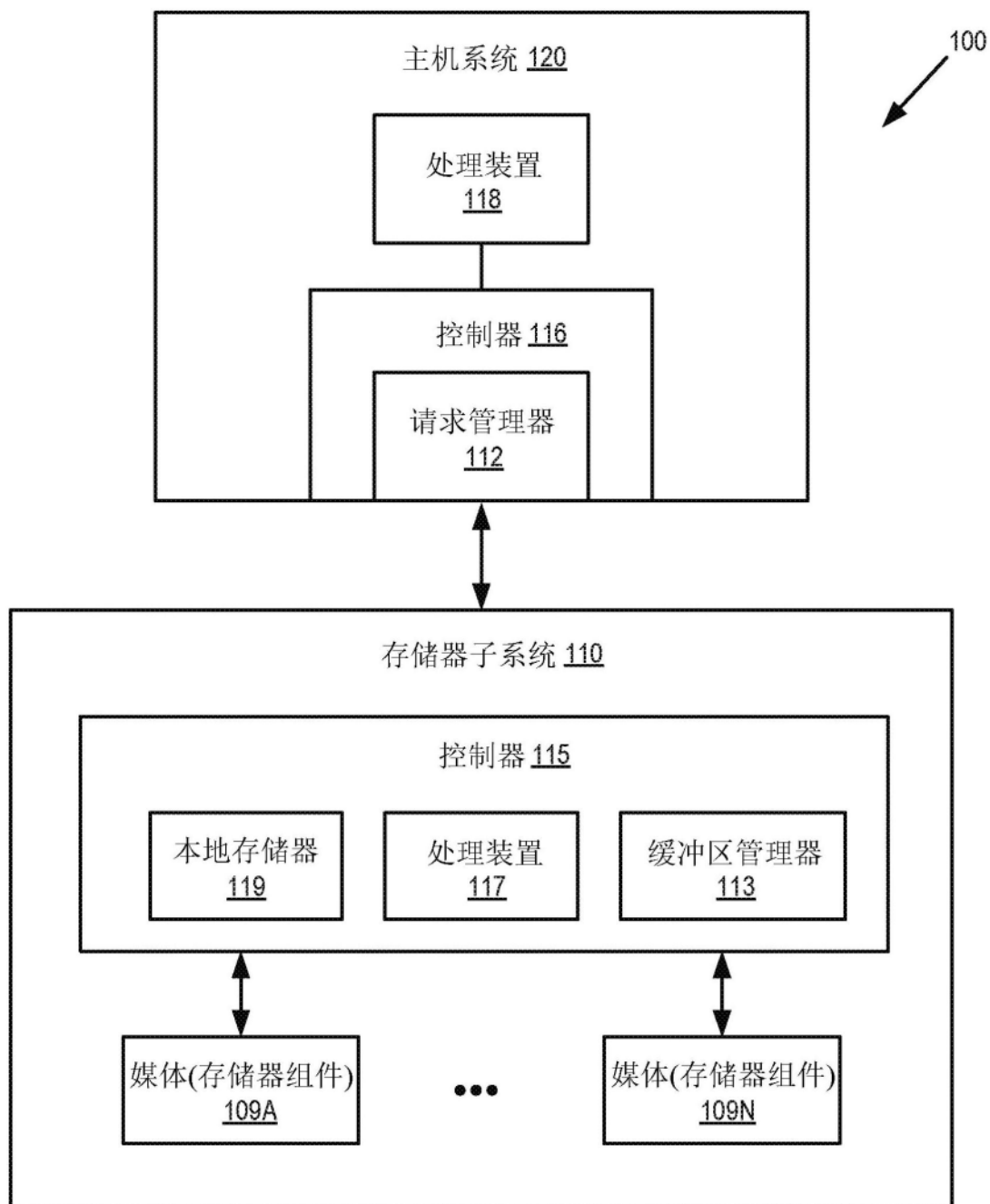


图1

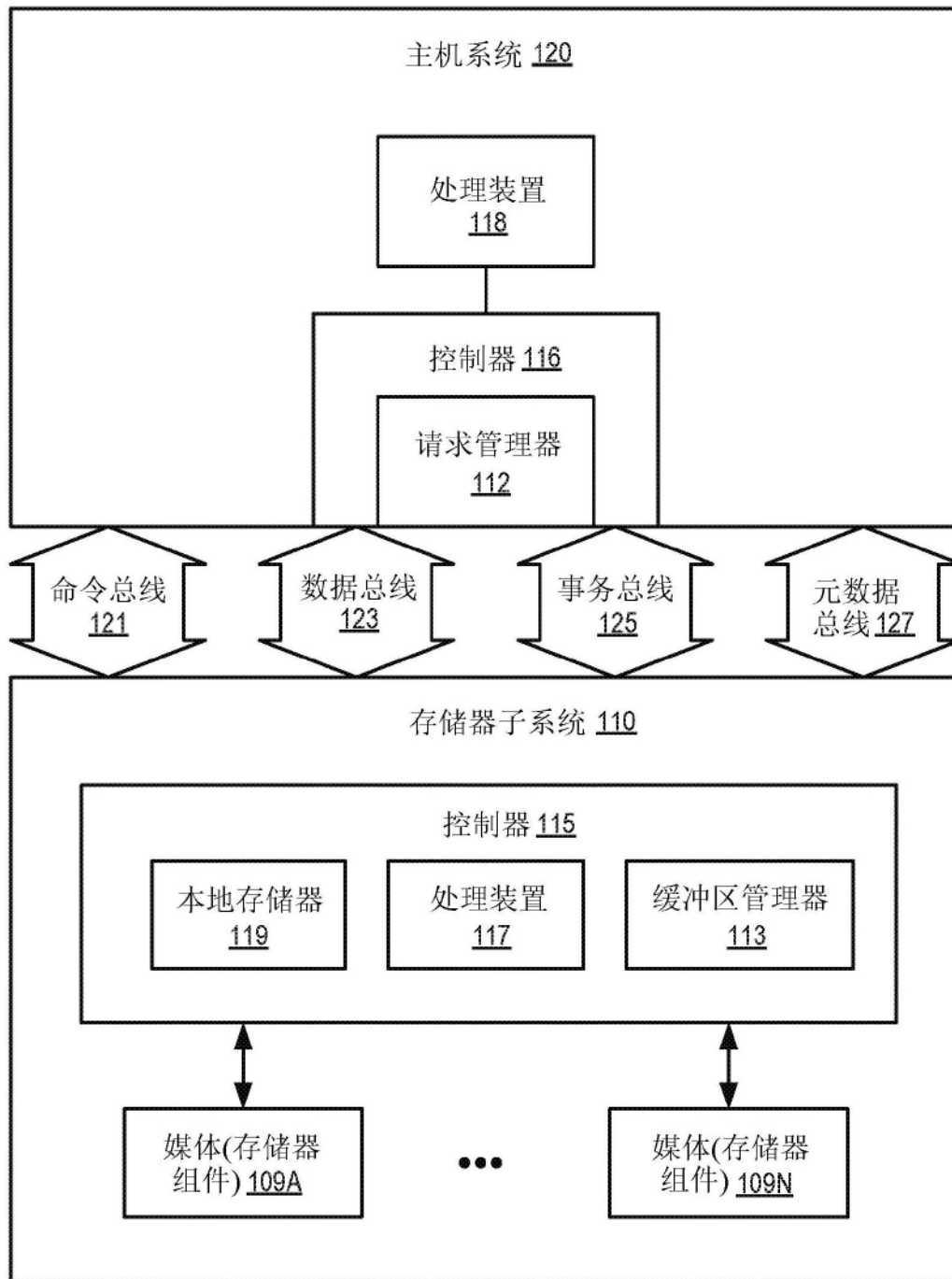


图2

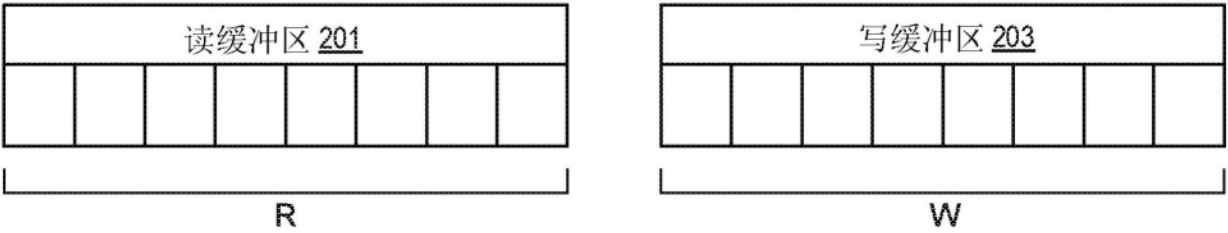


图3A



图3B



图3C

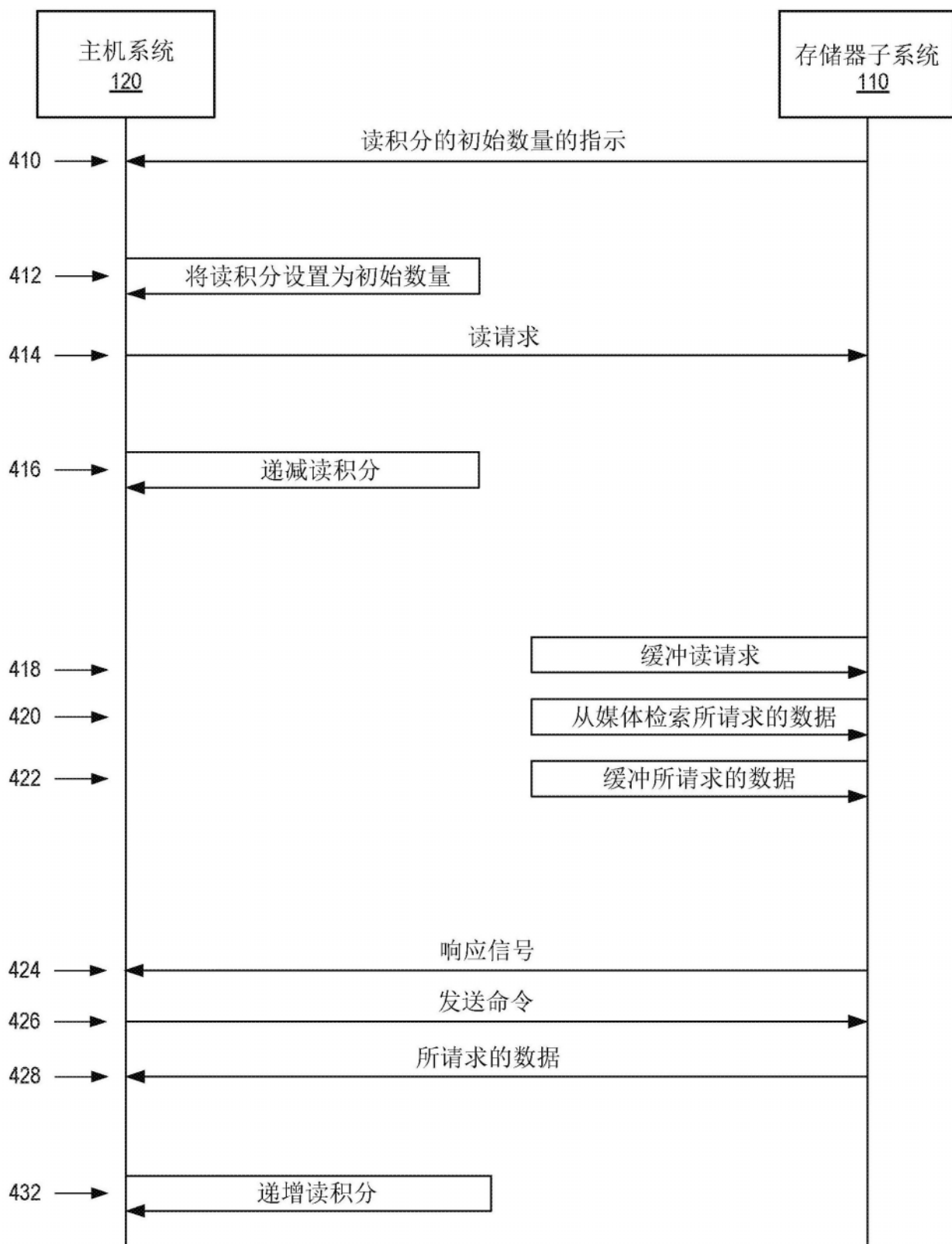


图4

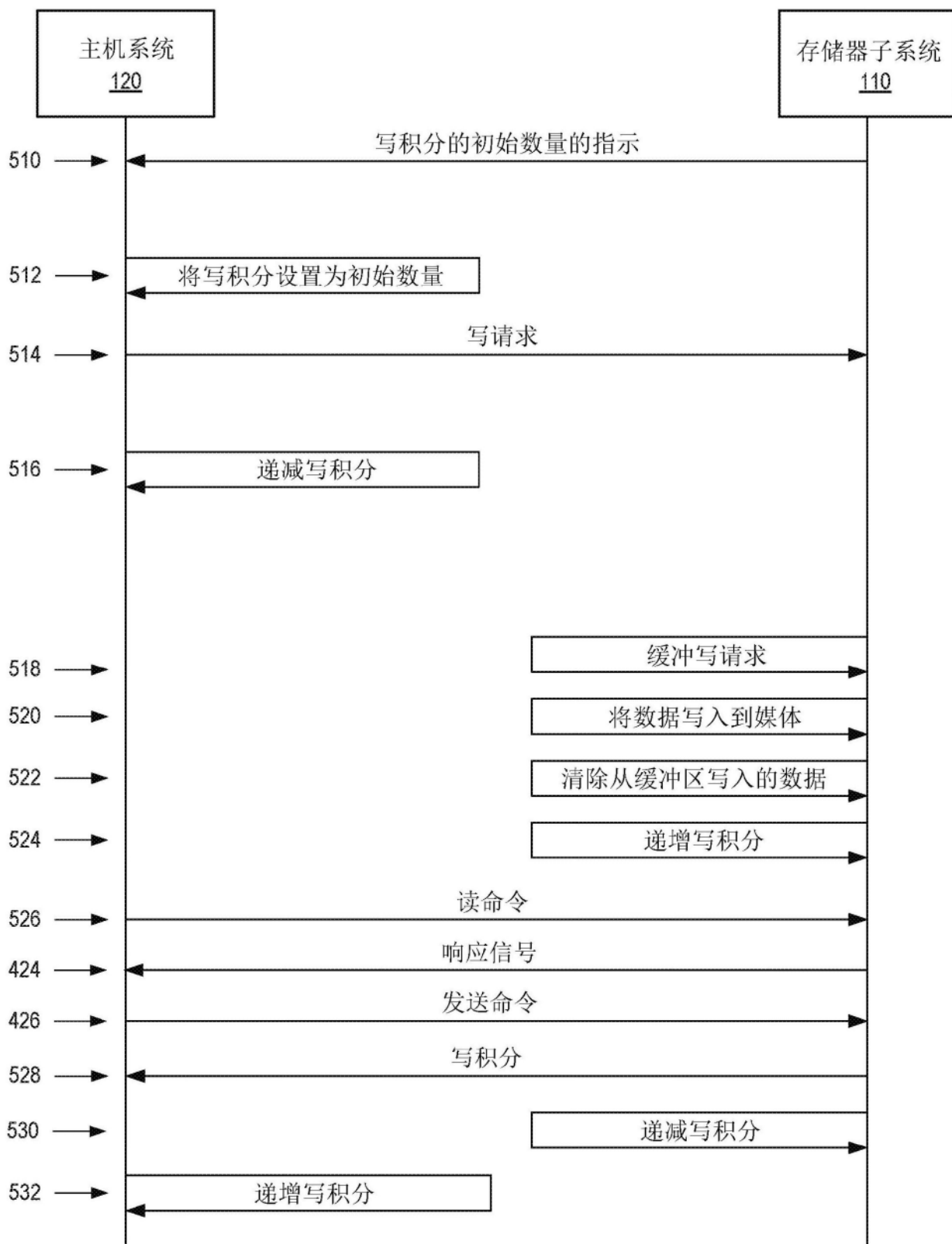


图5

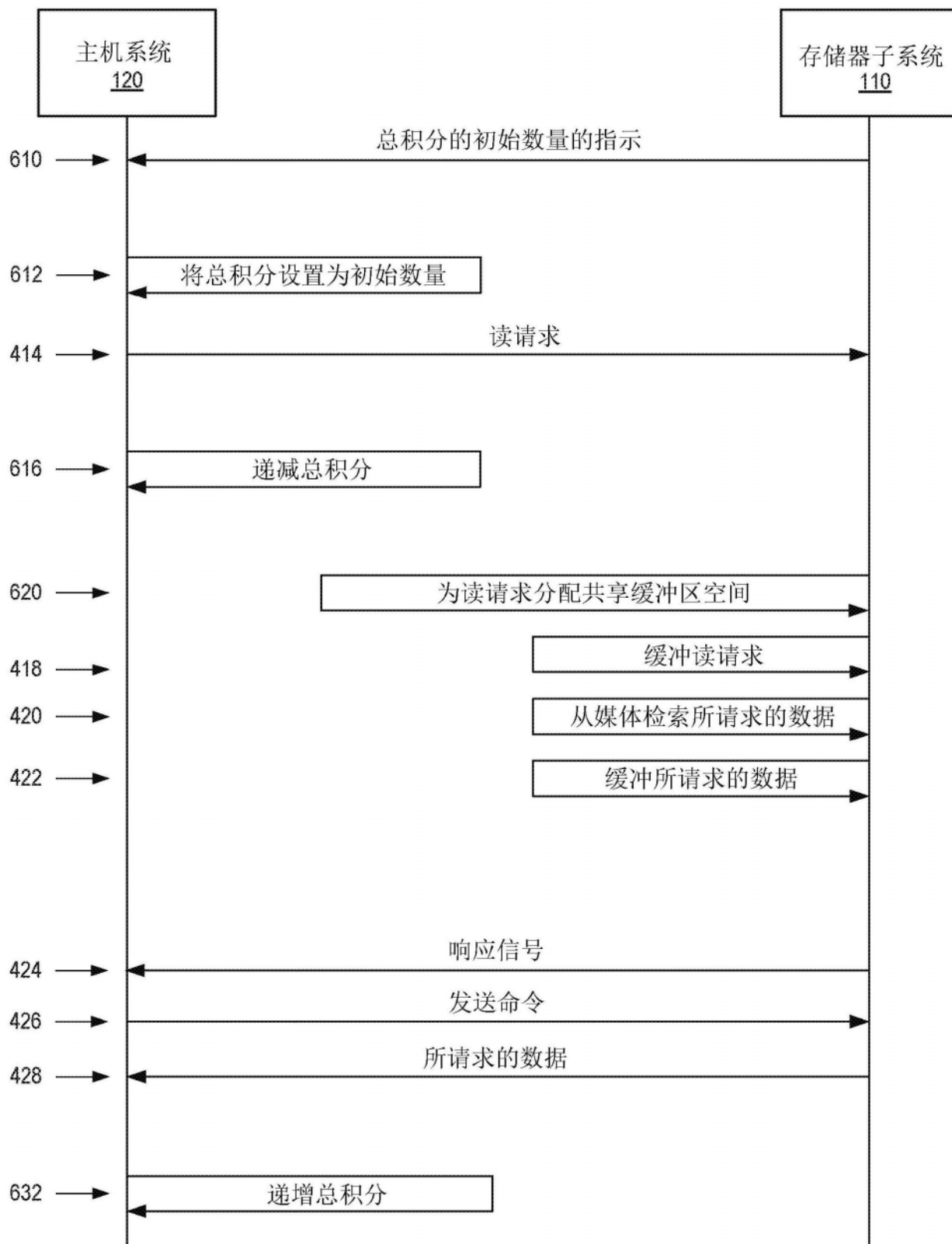


图6

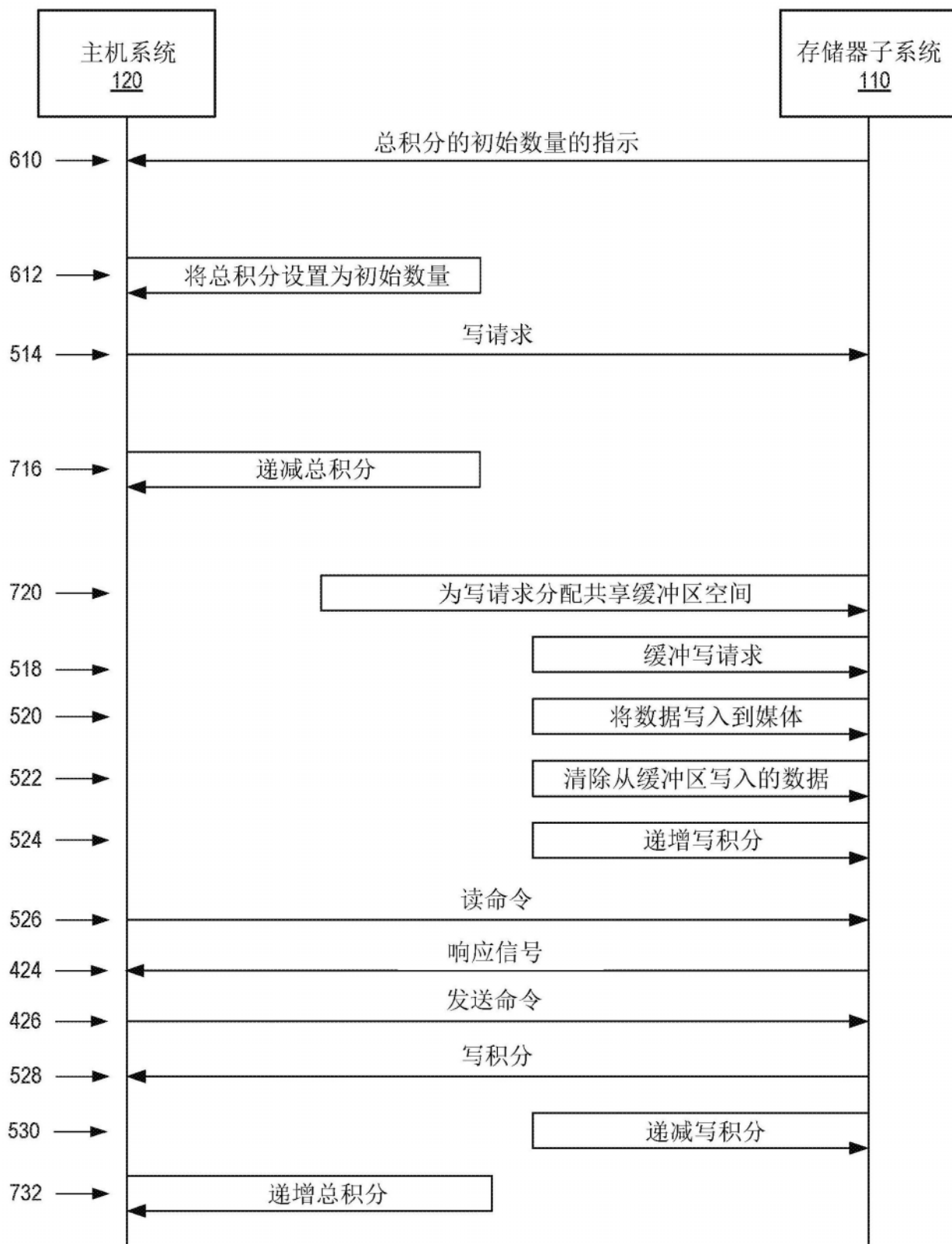


图7

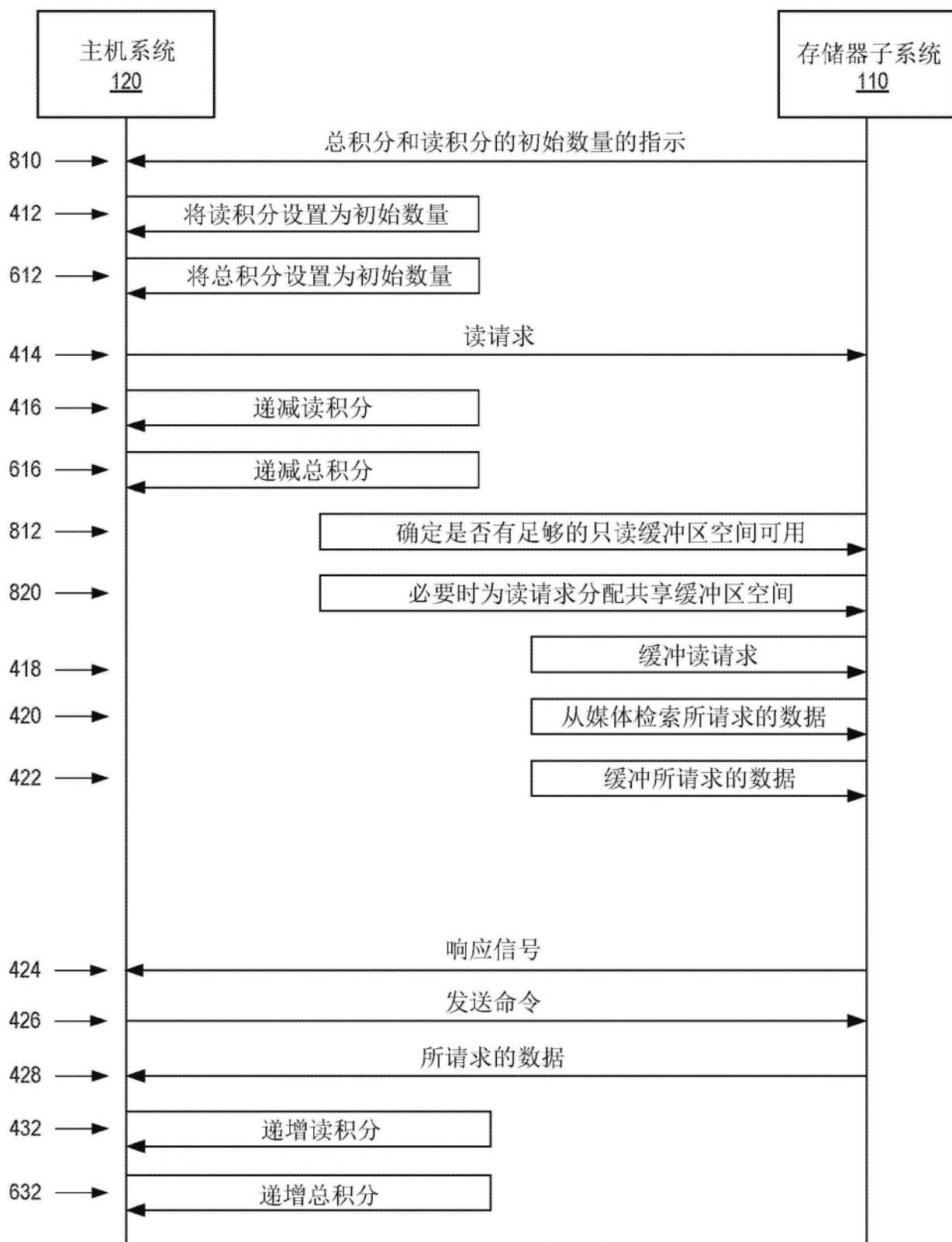


图8

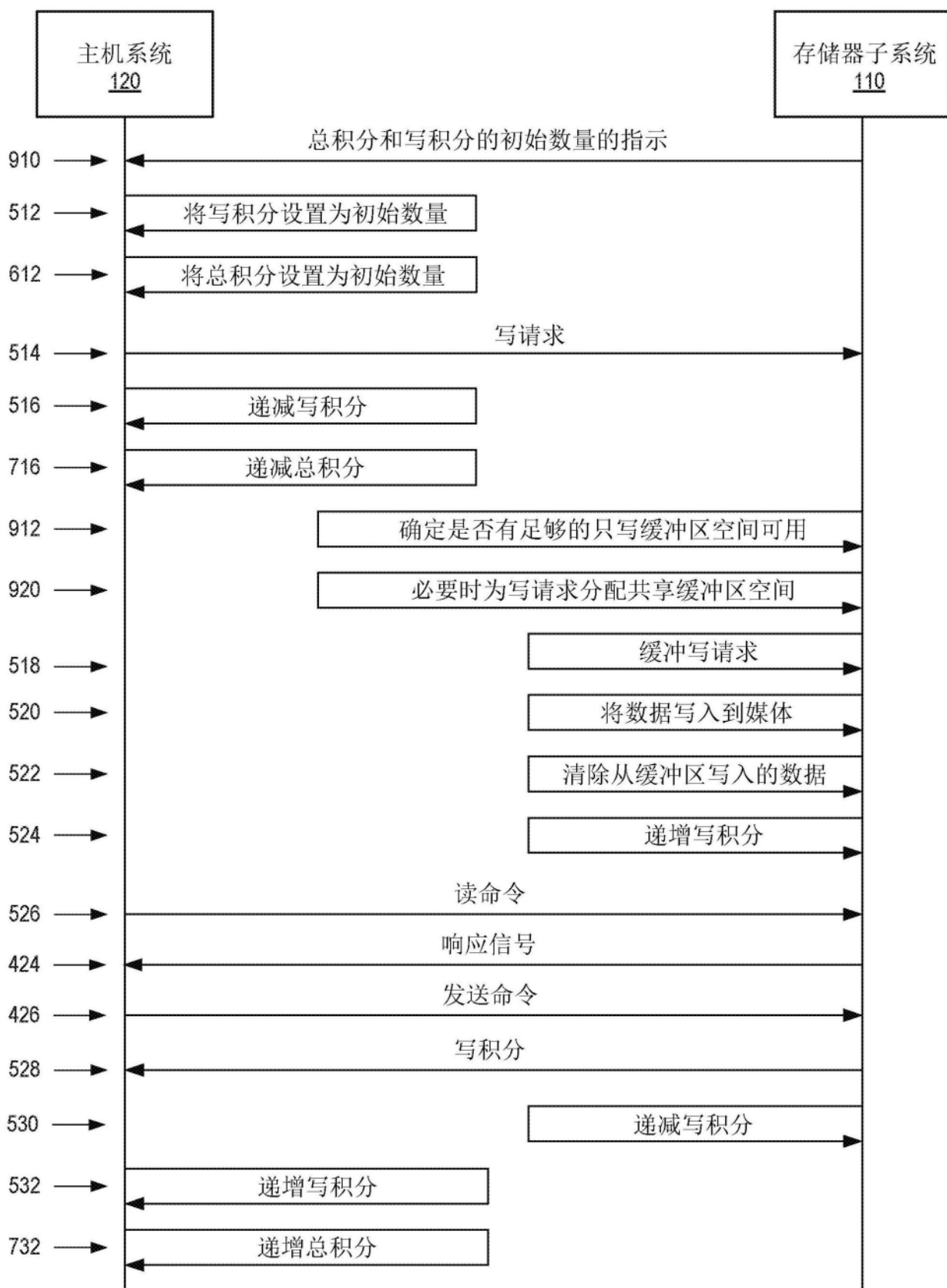


图9

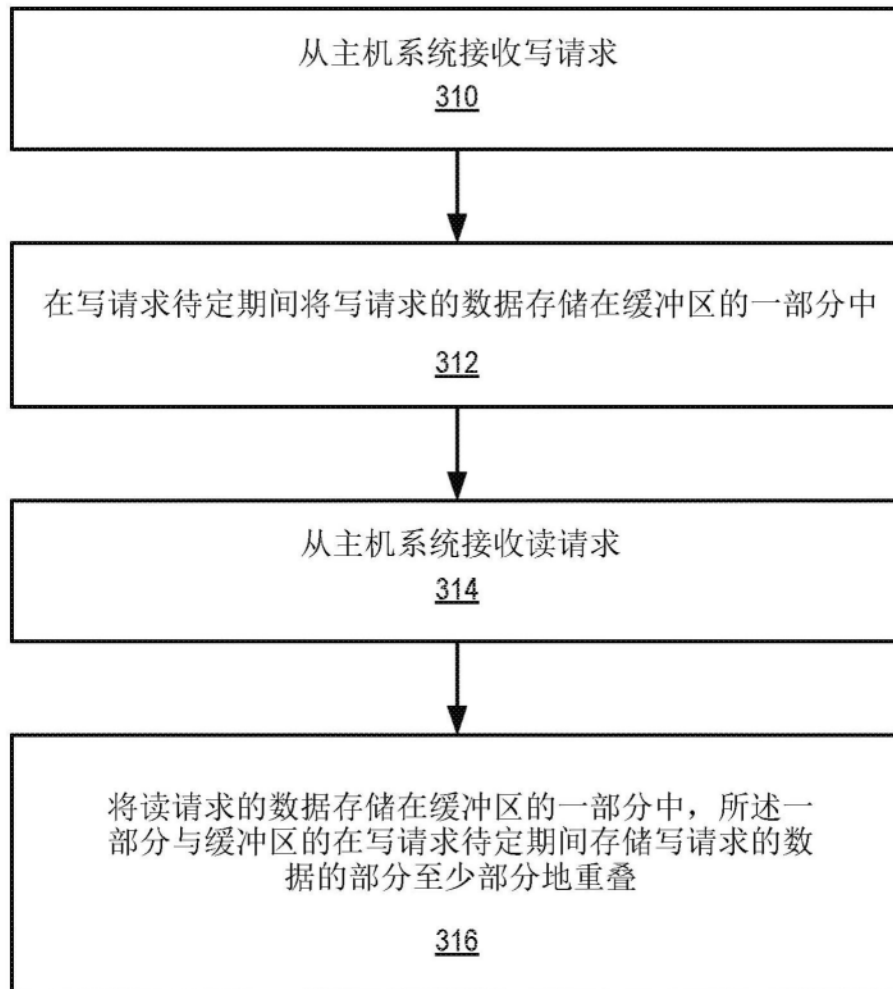


图10

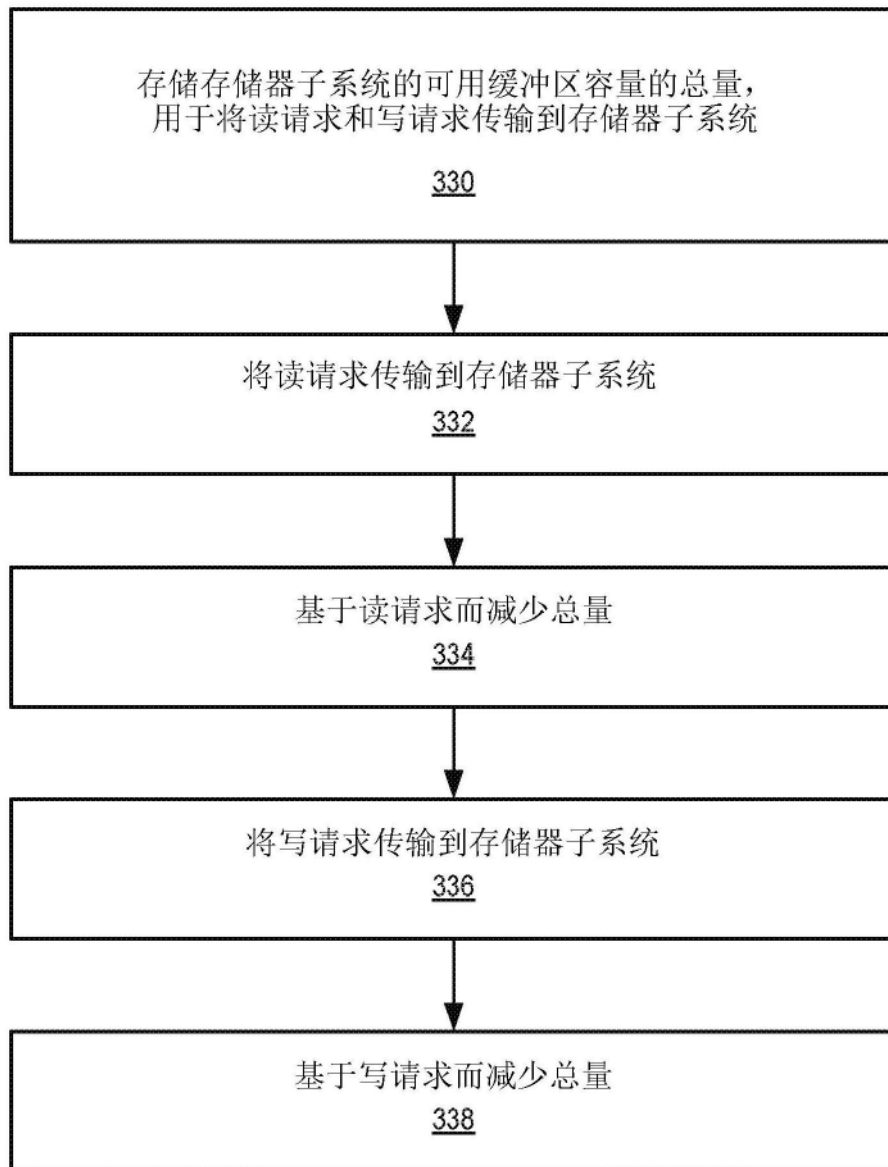


图11

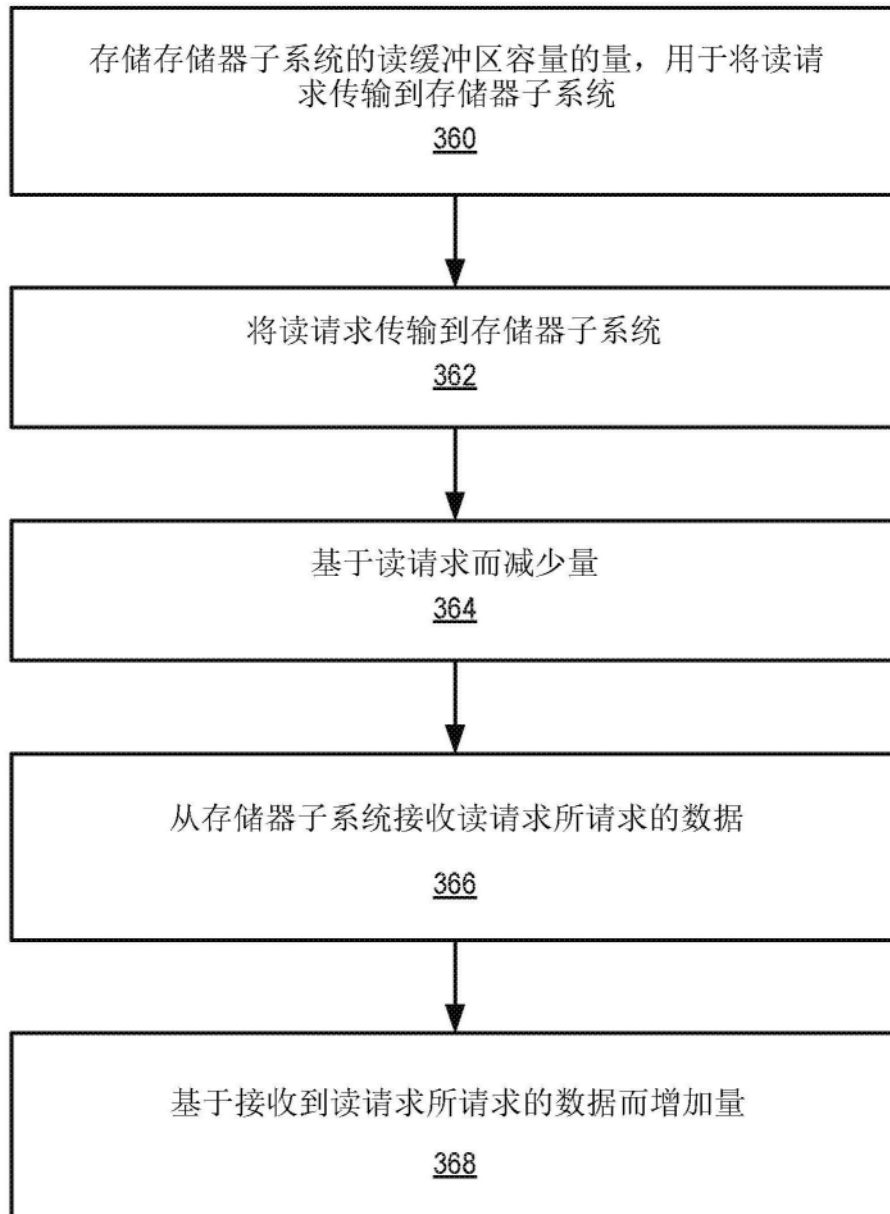


图12

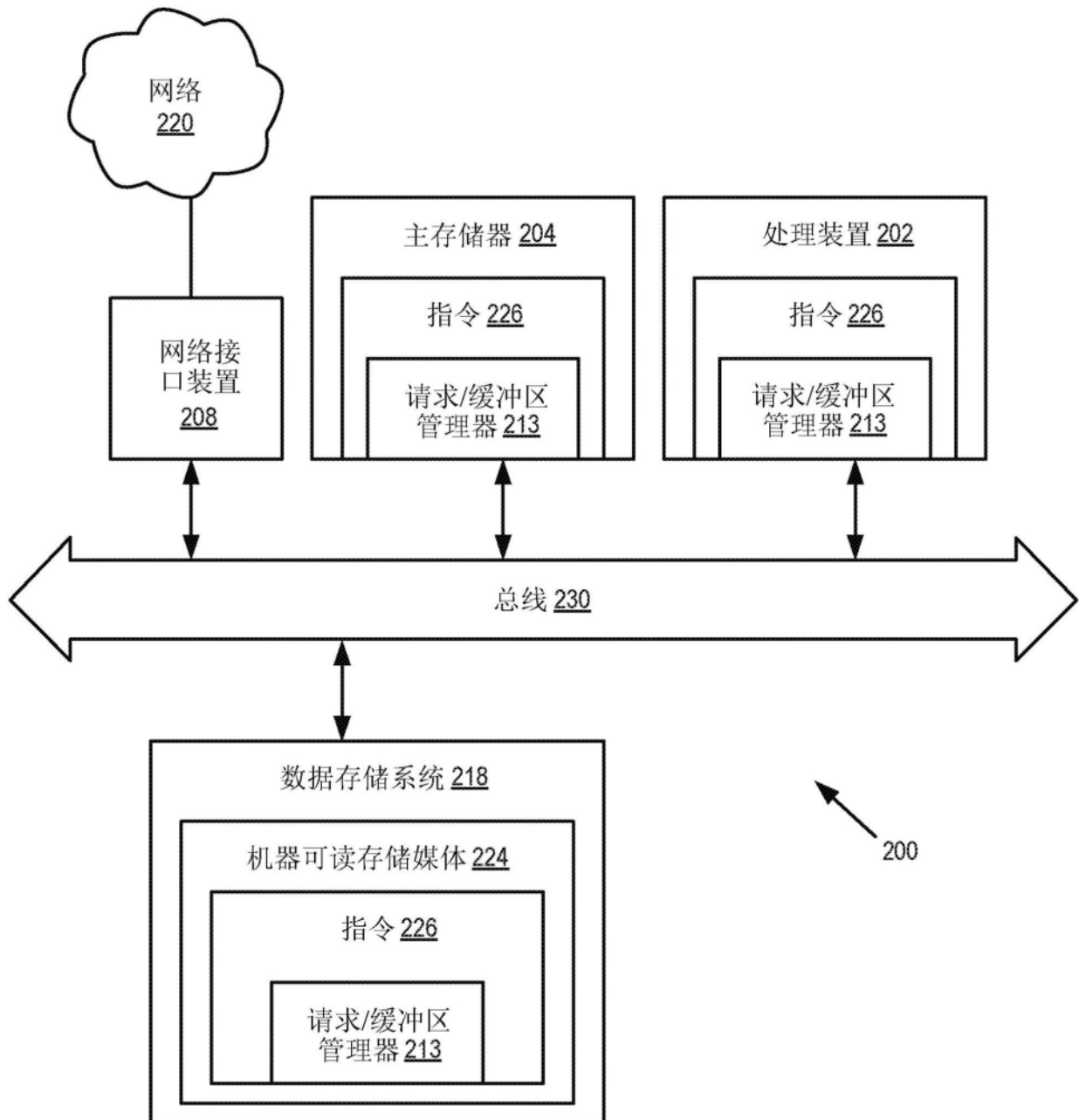


图13