

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号
WO 2013/170730 A1

(43) 国际公布日
2013年11月21日 (21.11.2013)

- (51) 国际专利分类号:
G06F 13/28 (2006.01)
- (21) 国际申请号: PCT/CN2013/075523
- (22) 国际申请日: 2013年5月11日 (11.05.2013)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201210147619.6 2012年5月12日 (12.05.2012) CN
- (71) 申请人: 北京忆恒创源科技有限公司 (MEMBLAZE TECHNOLOGY (BEIJING) CO., LTD.) [CN/CN]; 中国北京市海淀区上地信息路2号创业园D栋609, Beijing 100085 (CN)。
- (72) 发明人: 路向峰 (LU, Xiangfeng); 中国北京市海淀区上地信息路2号创业园D栋609, Beijing 100085 (CN)。

- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。
- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)。

[见续页]

(54) Title: DMA TRANSMISSION METHOD AND SYSTEM

(54) 发明名称: DMA 传输方法及系统

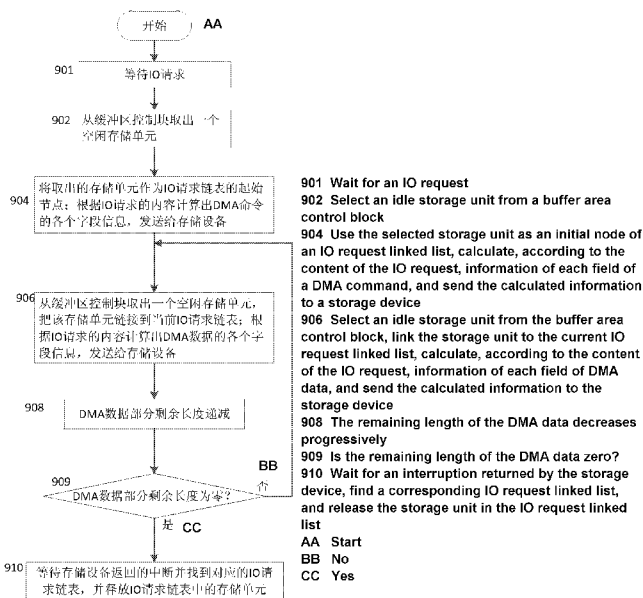


图 9A /FIG. 9A

(57) Abstract: Provided are a DMA transmission method and system. The DMA transmission method is a method for transmitting data between an information processing device and a storage device. The storage device comprises a buffer memory and a flash memory chip. The method comprises: receiving a first write request, the first write request comprising data to be written in and an address used for the flash memory chip of the storage device; distributing a first storage unit for the first write request in the information processing device; sending a first write command to the storage device, the first write command comprising data, the address used for the flash memory chip of the storage device and an address used for the buffer memory, and the address used for the buffer memory corresponding to the first storage unit; receiving a message from the storage device, the message indicating that the storage device finishes executing the write command; and releasing the first storage unit.

(57) 摘要: 提供了DMA传输方法及系统。一种在信息处理设备和存储设备之间进行数据传输的方法, 所述存储设备包括缓冲存储器和闪存芯片, 所述方法包括: 接收第一写请求, 所述第一写请求包括要写入的数据和用于所述存储设备的闪存芯片的地址; 为所述第一写请求在所述信息处理设备中分配第一存储单元; 向所述存储设备发送第一写命令, 所述第一写命令中包括数据、用于所述存储设备的闪存芯片的地址以及用于所述缓冲存储器的地址; 其中, 所述用于缓冲存储器的地址同所述第一存储单元相对应; 接收来自所述存储设备的消息, 所述消息指示所述存储设备对所述写命令已执行完成;

释放所述第一存储单元。

WO 2013/170730 A1

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

说明书

DMA 传输方法及系统

技术领域

本发明涉及固态存储设备 (Solid Storage Device, SSD), 更具体地, 本发明涉及将数据通过 DMA 传输写入固态存储设备的方法及装置。

背景技术

同机械式硬盘相类似, 固态存储设备 (SSD) 也是用于计算机系统的大容量、非易失性存储设备。固态存储设备一般以闪存 (Flash) 作为存储介质。在中国专利文献 CN102043689A 中公开了如图 13 所示的固态存储设备。如图 13 所示, 为目前一般的固态存储设备的功能框图。其中主要包括主机系统 1301 和固态存储设备 1302。其中, 固态存储设备 1302 包括接口模块 1303, 固态存储处理器 1304, 以及以 Flash 颗粒 1305 为单位组成的 Flash 阵列 1306。其中, 接口模块 1303 主要用于实现与主机系统一致的接口协议, 例如 SATA (Serial Advanced Technology Attachment, 串行高级技术附件)、USB (Universal Serial Bus, 通用串行总线)、PCI-E (Peripheral Component Interconnect Express, 快速外围组件互连)、SCSI (Small Computer System Interface, 小型计算机系统接口)、IDE (Integrated Drive Electronics, 集成驱动器电子) 等。通过接口模块 1303, 固态存储设备呈现给主机系统的是一个拥有一定逻辑空间的标准存储设备。固态存储处理器 1304 是整个存储设备的控制核心, 主要负责接口模块 1303 以及闪存阵列 1306 之间的控制信号及数据的传输、Flash 管理、主机逻辑地址到 Flash 物理地址的转换或映射、损耗均衡 (将逻辑地址映射到不同的物理地址从而防止单个 Flash 被过于集中地操作而提前失效)、坏块管理等。可由软件、硬件、固件或者其组合的多种方式实现固态存储处理器 1304。1305 为单个 Flash 颗粒, 多个 Flash 颗粒 1305 组成 Flash 阵列 1306。

为提高固态存储设备的读、写速度, 可以在固态存储设备中设置诸如 DRAM 或 SRAM 的随机访问存储器或其他类型的适于高速读/写操作的存储器, 作为同闪存读写数据时的高速缓冲存储器。在存储设备访问过程中, 作为一个例子, 计算机向存储设备发出 SCSI (小型计算机系统接口) 命令, 存储设备接收并处理 SCSI 命令, 依据 SCSI 命令所指示的操作执行相应的存储介质读写过程。在这一过程中, SCSI 命令并不直接操作高速缓冲存储器。即, 高速缓冲存储器对计算机或者用户是“透明”的。也有一些存储设备提供了高速缓冲存储器的“清洗”机制, 计算机或用户可使用预定的命令强制存储设备将高速缓冲存储器中的数据写入到非易失性存储介质 (例如, 磁盘或者闪存) 中。

然而, 对高速缓冲存储器的分配和管理将成为固态存储设备上的控制器的工作负担。而且当高速缓冲存储器被完全占用后, 如果固态存储设备接收到来自主机系统的新的访问请求, 则还需要执行对高速缓冲存储器的替换操作。这样既增加了控制器的复杂度, 又会让主机经历在读/写性能上的颠簸。

在主机与设备之间还可以进行 DMA (Direct Memory Access, 直接存储器访问) 传输。在中国专利文献 CN101221544A 中公开了执行 DMA 传输的方法和设备。

DMA 传输的一个典型过程是分散/收集 (Scatter/Gather) 操作。分散/收集操作中, 待传输的多个数据块存储在系统 (主机) 存储器的不连续的多个地址位置。处理器不需要为要从某一源移动到某一目的地的每个数据块对 DMA 控制器进行编程的操作。而是, 处理器在系统存储器中建立描述符表或描述符链接表。描述符表或描述符链接表中包括一组描述符。每个描述符都描述了数据块移动方向、源地址、目的地地址以及可选的传输的字节数。在一个描述符中不包括传输字节数的情况下, 可通过 DMA 方式传输约定长度的数据。

发明内容

因而, 分担存储设备上的控制器的工作负担是有益的。通过将存储设备的缓冲存储器有关的维护工作转移给主机, 既减轻了存储设备的控制器的负载, 也向主机提供了更灵活控制存储设备的能力。

在本发明的第一实施例中, 提供了一种在信息处理设备和存储设备之间进行 DMA 传输的方法, 所述存储设备包括缓冲存储器和闪存芯片, 所述方法包括: 接收第一 IO 请求; 为所述第一 IO 请求分配第一存储单元与第二存储单元; 向所述存储设备发送所述第一 DMA 描述符, 所述第一 DMA 描述符包括 DMA 主机地址、用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址与第二用于缓冲存储器的地址, 其中, 所述第一用于缓冲存储器的地址同所述第一存储单元相对应, 所述第二用于缓冲存储器的地址同所述第二存储单元相对应; 在所述存储设备和所述信息处理设备之间依据所述第一 DMA 描述符进行 DMA 传输; 接收来自所述存储设备的消息, 所述消息指示所述存储设备对第一 DMA 描述符已执行完成; 释放所述第一存储单元与所述第二存储单元。

根据本发明的第一实施例, 其中分配第一存储单元与第二存储单元进一步包括, 在所述第二存储单元中存储指向所述第一存储单元的地址。

在本发明的第二实施例中提供了一种在信息处理设备和存储设备之间进行 DMA 传输的方法, 所述存储设备包括缓冲存储器和闪存芯片, 所述方法包括: 接收第一 IO 请求; 为所述第一 IO 请求分配第一存储单元与第二存储单元; 向所述存储设备发送所述第一 DMA 描述符, 所述第一 DMA 描述符包括第一 DMA 描述符命令和第一 DMA 描述符数据, 所述第一 DMA 描述符命令包括用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址, 所述第一 DMA 描述符数据包括 DMA 主机地址以及第二用于缓冲存储器的地址; 其中, 所述第一用于缓冲存储器的地址同所述第一存储单元相对应, 所述第二用于缓冲存储器的地址同所述第二存储单元相对应; 在所述存储设备和所述信息处理设备之间依据所述 DMA 主机地址和所述第二用于缓冲存储器的地址进行 DMA 传输; 接收来自所述存储设备的消息, 所述消息指示所述存储设备对第一 DMA 描述符已执行完成; 释放所述第一存储单元与所述第二存储单元。

根据本发明的第一或第二实施例, 其中, 所述第一 IO 请求指示将所述 DMA 主机地址处的数据写入到用于所述存储设备的闪存芯片的地址处。

根据本发明的第一或第二实施例, 还包括将空闲的存储单元链接在一起形成空闲存储单元资源池, 在分配第一存储单元与第二存储单元时, 从所述空闲存储单元资源池中取出空闲的存储单元, 在释放所述第一存储单元与所述第二存储单元

时,基于所述所述第一存储单元与所述第二存储单元插入到所述空闲存储单元资源池。

根据本发明的第一或第二实施例,还包括将空闲的存储单元链接在一起形成空闲存储单元资源池,在分配第一存储单元与第二存储单元时,若所述空闲存储单元资源池为空,则等待一个或多个存储单元被插入到所述空闲存储单元资源池。

在本发明的第三实施例中,提供了一种在信息处理设备和存储设备之间进行DMA传输的方法,所述存储设备包括缓冲存储器和闪存芯片,所述方法包括:接收第一IO请求;为所述第一IO请求分配第一存储单元、第二存储单元与第三存储单元;向所述存储设备发送所述第一DMA描述符,所述第一DMA描述符包括第一DMA描述符命令、第一DMA描述符数据和第二DMA描述符数据,所述第一DMA描述符命令包括用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址,所述第一DMA描述符数据包括第一DMA主机地址以及第二用于缓冲存储器的地址,所述第二DMA描述符数据包括第二DMA主机地址以及第三用于缓冲存储器的地址;其中,所述第一用于缓冲存储器的地址同所述第一存储单元相对应,所述第二用于缓冲存储器的地址同所述第二存储单元相对应,所述第三用于缓冲存储器的地址同所述第三存储单元相对应;在所述存储设备和所述信息处理设备之间依据所述第一DMA主机地址和所述第二用于缓冲存储器的地址进行DMA传输,以及依据所述第二DMA主机地址和所述第三用于缓冲存储器的地址进行DMA传输;接收来自所述存储设备的消息,所述消息指示所述存储设备对第一DMA描述符已执行完成;释放所述第一存储单元、所述第二存储单元与所述第三存储单元。

根据本发明的第三实施例,所述第一IO请求指示将所述第一DMA主机地址处的数以及所述第二DMA主机地址处的数据写入到用于所述存储设备的闪存芯片的地址处。根据本发明的第三实施例,其中分配第一存储单元、第二存储单元与第三存储单元后,将所述第一存储单元、所述第二存储单元和所述第三存储单元链接为循环链表。根据本发明的第三实施例,所述来自所述存储设备的消息中包括指示所述第一存储单元、所述第二存储单元和所述第三存储单元之一的内容。

根据本发明的第一、第二和第三实施例,其中分配所述第一存储单元后,在所述第一存储单元中存储用于标识所述第一IO请求的指针。

在本发明的第四实施例中提供了一种在信息处理设备和存储设备之间进行数据传输的方法,所述存储设备包括缓冲存储器和闪存芯片,所述方法包括:接收第一写请求,所述第一写请求包括要写入的数据和用于所述存储设备的闪存芯片的地址;为所述第一写请求在所述信息处理设备中分配第一存储单元;向所述存储设备发送第一写命令,所述第一写命令中包括数据、用于所述存储设备的闪存芯片的地址以及用于所述缓冲存储器的地址;其中,所述用于缓冲存储器的地址同所述第一存储单元相对应;接收来自所述存储设备的消息,所述消息指示所述存储设备对所述第一写命令已执行完成;释放所述第一存储单元。

根据本发明的第四实施例,其中还为所述第一写请求分配第二存储单元,并在所述第二存储单元中存储指向所述第一存储单元的指针;基于所述消息,释放所述第一存储单元与所述第二存储单元。

在本发明的第五实施例中,提供了一种在信息处理设备和存储设备之间进行 DMA 传输的装置,所述存储设备包括缓冲存储器和闪存芯片,所述装置包括:用于接收第一 IO 请求的模块;用于为所述第一 IO 请求分配第一存储单元与第二存储单元的模块;用于向所述存储设备发送所述第一 DMA 描述符的模块,所述第一 DMA 描述符包括 DMA 主机地址、用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址与第二用于缓冲存储器的地址,其中,所述第一用于缓冲存储器的地址同所述第一存储单元相对应,所述第二用于缓冲存储器的地址同所述第二存储单元相对应;用于在所述存储设备和所述信息处理设备之间依据所述第一 DMA 描述符进行 DMA 传输的模块;用于接收来自所述存储设备的消息的模块,所述消息指示所述存储设备对第一 DMA 描述符已执行完成;用于释放所述第一存储单元与所述第二存储单元的模块。

根据本发明的第五实施例,还包括用于在分配第一存储单元、第二存储单元与第三存储单元后,将所述第一存储单元、所述第二存储单元和所述第三存储单元链接为循环链表的装置。

根据本发明的第五实施例,还包括:用于将空闲的存储单元链接在一起形成空闲存储单元资源池的装置;在分配第一存储单元与第二存储单元时,从所述空闲存储单元资源池中取出空闲的存储单元,在释放所述第一存储单元与所述第二存储单元时,基于所述所述第一存储单元与所述第二存储单元插入到所述空闲存储单元资源池。根据本发明的第五实施例,还包括:用于将空闲的存储单元链接在一起形成空闲存储单元资源池的装置,在分配第一存储单元与第二存储单元时,若所述空闲存储单元资源池为空,则等待一个或多个存储单元被插入到所述空闲存储单元资源池。

本发明的第六实施例提供了一种在信息处理设备和存储设备之间进行 DMA 传输的装置,所述存储设备包括缓冲存储器和闪存芯片,所述装置包括:用于接收第一 IO 请求的模块;用于为所述第一 IO 请求分配第一存储单元与第二存储单元的模块;用于向所述存储设备发送所述第一 DMA 描述符的模块,所述第一 DMA 描述符包括第一 DMA 描述符命令和第一 DMA 描述符数据,所述第一 DMA 描述符命令包括用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址,所述第一 DMA 描述符数据包括 DMA 主机地址以及第二用于缓冲存储器的地址;其中,所述第一用于缓冲存储器的地址同所述第一存储单元相对应,所述第二用于缓冲存储器的地址同所述第二存储单元相对应;用于在所述存储设备和所述信息处理设备之间依据所述 DMA 主机地址和所述第二用于缓冲存储器的地址进行 DMA 传输的模块;用于接收来自所述存储设备的消息的模块,所述消息指示所述存储设备对第一 DMA 描述符已执行完成;用于释放所述第一存储单元与所述第二存储单元的模块。

本发明的第七实施例提供了一种在信息处理设备和存储设备之间进行数据传输的装置,所述存储设备包括缓冲存储器和闪存芯片,所述装置包括:用于接收第一写请求的装置,所述第一写请求包括要写入的数据和用于所述存储设备的闪存芯片的地址;用于为所述第一写请求在所述信息处理设备中分配第一存储单元的装置;用于向所述存储设备发送第一写命令的装置,所述第一写命令中包括数据、用

于所述存储设备的闪存芯片的地址以及用于所述缓冲存储器的地址；其中，所述用于缓冲存储器的地址同所述第一存储单元相对应；用于接收来自所述存储设备的消息的装置，所述消息指示所述存储设备对所述写命令已执行完成；用于释放所述第一存储单元的装置。

根据第七实施例，还包括用于为所述第一写请求分配第二存储单元，并在所述第二存储单元中存储指向所述第一存储单元的指针的装置；以及用于基于所述消息，释放所述第一存储单元与所述第二存储单元的装置。根据本发明的第五实施例、第六实施例和第七实施例，还包括用于在分配所述第一存储单元后，在所述第一存储单元中存储用于标识所述第一 IO 请求的指针的装置。

附图说明

当连同附图阅读时，通过参考后面对示出性的实施例的详细描述，将最佳地理解本发明以及优选的使用模式和其进一步的目的和优点，其中附图包括：

图 1 是根据本发明实施例的存储设备的结构框图；

图 2A、2B 是根据本发明实施例的写入命令的示意图；

图 3 是根据本发明实施例的存储设备执行写入命令的方法的流程图；

图 4 是根据本发明实施例的主机的示意图；

图 5 是根据本发明实施例的主机执行写入操作的流程图；

图 6 是根据本发明实施例的第二写入命令的示意图；

图 7A、7B 是根据本发明实施例的存储设备执行第二写入命令的流程图；

图 7C 中展示了用于实施图 7A、7B 中的执行第二写入命令的存储设备的硬件方框图；

图 8 是根据本发明实施例的主机的软件方框图；

图 9A 是根据本发明实施例的主机创建并执行第二写入命令的流程图；

图 9B 是根据本发明另一实施例的主机创建并执行第二写入命令的流程图；

图 10A 是根据本发明的一实施例的在存储设备的缓冲存储器中创建链表的流程图；

图 10B 是根据本发明的一实施例的在存储设备的缓冲存储器中创建链表的流程图；

图 10C 是根据本发明的一实施例的存储设备利用在缓冲存储器中创建的链表执行 DMA 描述符的流程图；

图 11A-11F 展示出了与图 10B、图 10C 相关联的缓冲存储器的多种状态；

图 12 是根据本发明的又一实施例的存储设备的硬件方框图；

图 13 是根据现有技术的存储设备的方框图。

具体实施方式

图 1 是根据本发明实施例的存储设备的结构框图。如图 1 所示的实施例包括主机 101 以及同主机 101 相耦合的存储设备 102。主机 101 同存储设备 102 之间可通过多种方式相耦合，耦合方式包括但不限于通过例如 SATA、IDE、USB、PCIE、SCSI、以太网、光纤通道、无线网络等连接主机 101 与存储设备 102。主机 101 可以是能够通过上述方式同存储设备相通信的信息处理设备，例如，个人计算

机、平板电脑、服务器、便携式计算机、网络交换机、路由器、蜂窝电话、个人数字助理等。存储设备 102 包括主机接口 103、控制电路 104、一个或多个闪存芯片 105 以及缓冲存储器 106。主机接口 103 可适配于通过例如 SATA、IDE、USB、PCI-E、SCSI、以太网、光纤通道等方式与主机 101 交换数据。控制电路 104 用于控制在主机接口 103、闪存芯片 105 以及缓冲存储器 106 之间的数据传输，还用于闪存管理、主机逻辑地址到闪存物理地址映射、擦除均衡、坏块管理等。可通过软件、硬件、固件或其组合的多种方式实现控制电路 104。控制电路 104 可以是 FPGA (Field-programmable gate array, 现场可编程门阵列)、ASIC (Application Specific Integrated Circuit, 应用专用集成电路) 或者其组合的形式。控制电路 104 也可以包括处理器或者控制器。

根据本发明的一个实施例，主机 101 向存储设备 102 发出读出命令或者写入命令。控制电路 104 经由主机接口 103 接收到该读出命令或写入命令。在附图 2 中详细描述了作为例子的第一写入命令 200。

参看图 2A、2B，图 2A 是根据本发明实施例的写入命令的示意图。写入命令 200 包括字段 201、202、203 和 204。字段 201 指示该命令为写入命令，字段 202 为闪存地址，字段 203 为数据，该写入命令 200 指示存储设备 102 将数据字段 203 中的数据基于由字段 202 所指示的闪存地址写入闪存芯片 105。字段 204 为缓冲存储器地址，存储设备 102 接收到该写入命令 200 时，先将数据字段 203 中的数据基于写入由字段 204 所指示缓冲存储器地址写入缓冲存储器 106，再将数据字段 203 中的数据基于由字段 202 所指示的闪存地址写入闪存芯片 105。在一个例子中，存储设备 102 将数据字段 203 中的数据基于写入由字段 204 所指示缓冲存储器地址写入缓冲存储器 106 后，再从缓冲存储器 106 中读出该数据，继而将该数据写入到闪存芯片 105。将数据写入到缓冲存储器 106 的操作，同将存储器 106 中的另一数据读出并写入到闪存芯片 105 的操作可以并发执行，从而使得主机 101 向存储设备 102 的写入操作的并发性能得到提升，并且不会显著增加控制电路 104 的复杂度，因为控制电路 104 无需处理缓冲存储器 106 的空间分配任务。在一个实施例中，字段 204 中可以是缓冲存储器 106 的完整地址，而在另一个实施例中，字段 204 是相对某一基地址的偏移值。在一个实施例中，字段 203 中携带要写入闪存 105 的数据。而在另一个实施例中，字段 203 中可携带一个指针，该指针指向要写入存储设备 102 的数据，而该数据可以存储在主机 101 的存储器中，在此情况下，存储设备 102 通过随后的 DMA 传输过程从主机 101 获得该数据。在依然另一个实施例中，字段 203 可携带一个指针，该指针指向要写入存储设备 102 的数据，而该数据可以存储在存储设备 102 的缓冲存储器中。字段 202 可以是要将数据写入的闪存芯片 105 的物理地址或逻辑地址。字段 202 也可以是一个指针，指向缓冲存储器 106，在其中存储有用于闪存芯片 105 的物理地址或逻辑地址。逻辑地址到物理地址的转换过程，可以通过查找地址映射表的方式实现。

所属领域技术人员将容易意识到，写入命令可以具有多种具体编码方式和字段顺序。例如，参看图 2B，指示写入命令 210 的类型是写操作的字段 214 可以在写入命令 210 的末尾或者其他位置。而在字段 211 中携带缓冲存储器地址，在字段 212 中携带数据或指向数据的存放位置的指针。在字段 213 中存放闪存地址，或者

指向存放闪存地址的指针，闪存地址可以是逻辑地址或物理地址。

图 3 是根据本发明实施例的存储设备执行写入命令的方法的流程图。在步骤 301，存储设备 102 从主机 101 接收到写入命令 200。存储设备 102 的控制电路 104 通过主机接口 103 接收到写入命令 200 后，提取出包含于写入命令 200 中的用于指示操作类型为写操作的字段 201，用于指示要写入的闪存的地址的字段 202，用于指示要写入的数据的字段 203，以及用于指示缓冲存储器 106 的地址的字段 204。在步骤 302，响应于该写入命令 200，控制电路 104 基于字段 204 得到用于缓冲存储器 106 的地址，并基于字段 203 得到要写入的数据，以及将要写入的数据写入到缓冲存储器 106 中由字段 204 所指示的位置。当将数据写入到缓冲存储器 106 中后，存储设备 102 可向主机发送消息以指示写入命令 200 的执行完成，虽然数据此时尚未被实际写入到闪存芯片 105 之中。以此方式，在主机 101 看来，在步骤 302 执行完成后，写入命令 200 已经执行完成，从而提升了存储设备 102 的执行写入命令 200 的性能。向主机发送的消息可以包含于存储设备 102 向主机 101 所发送的中断请求之中，也可以基于主机 101 与存储设备 102 之间的耦合方式（SATA、IDE、USB、PCIE、SCSI、以太网、光纤通道、无线通信网络等）而选择适当的其他消息发送方式。在将要写入的数据写入到缓冲存储器 106 之后，在控制电路 104 的控制下，将由字段 203 所指示的要写入的数据，基于由字段 202 所指示的闪存地址，写入到闪存芯片 105 中（步骤 303）。当将数据写入到闪存芯片 105 中后，存储设备 102 也可向主机发送消息以指示写入命令 200 的执行完成，特别地，此时主机可以在写入命令 200 中再次指定将数据写入该用于指示缓冲存储器 106 的地址，而不会因对该地址处的数据的重写而导致数据错误。在一个例子中，如果字段 202 所指示的是用于闪存芯片 105 的逻辑地址，则将该逻辑地址转换为用于闪存芯片 105 的物理地址。逻辑地址到物理地址的转换方式是所属领域技术人员所了解的。在一个例子中，在步骤 303，从缓冲存储器 106 中重新取得所写入的数据，并将该数据写入到闪存芯片 105。所属领域技术人员将意识到，在控制电路 104 的控制下，步骤 302 中将数据写入到缓冲存储器 106 的操作，与步骤 303 中将数据写入到闪存存储器 105 的操作，可以并行执行。这样，在存储设备 102 中可以同时处理多个写入命令，其中，在一个时刻，基于一个写入命令，控制电路 104 将第一数据写入到缓冲存储器 106 中；而基于另一个写入命令，控制电路 104 将存在于缓冲存储器 106 中的第二数据写入到闪存芯片 105 中。缓冲存储器 106 可以是双端口存储器，使得在经由第一端口向缓冲存储器 106 写入第一数据的同时，可以从第二端口从缓冲存储器 106 读出第二数据。所属领域技术人员可意识到缓冲存储器 106 的其他实施方式，以支持对多份数据的同时读出和/或写入操作。

通过在写入命令中携带用于指示缓冲存储器 106 的地址的字段 204，将维护缓冲存储器 106 的工作从控制电路 104 移除了，并且使主机 101 拥有了更灵活控制存储设备 102 的能力。

图 4 是根据本发明实施例的主机的示意图。图 4 是示出了主机 400 的软件组成的方框图。主机 400 可以是个人计算机、服务器计算机或者其他具有计算能力的设备。主机 400 包括一个或多个用户应用程序 401、402 和 403，以及操作系统 404。操作系统 404 中具有存储设备驱动程序 405。在根据本发明的实施例中，驱动

程序 405 中包括缓冲区控制块 406，用以在主机 400 中控制存储设备 102 的缓冲存储器 106。缓冲区控制块 406 由多个存储单元 (411、412... ..41n) 组成，缓冲区控制块 406 中的每个存储单元 (411、412... ..41n) 对应于缓冲存储器 106 中的一个存储单元，并记录缓冲存储器 106 中的对应存储单元的工作状态。在一个实施例中，缓冲区控制块 406 中的存储单元 (411、412... ..41n) 的每一个，记录缓冲存储器 106 中的对应存储单元是空闲的还是已经被占用。在进一步的一个实施例中，在发送给存储设备 102 的一个读/写命令中涉及缓冲存储器中的多个存储单元，例如 2 个。在此情况下，将缓冲区控制块 406 中的 2 个存储单元 (411、412) 关联在一起，在存储单元 411、412 中还记录存储单元 411 与 412 之间的这种关联关系，例如，在存储单元 411 中记录指向存储单元 412 的一个或多个指针。在依然进一步的实施例中，在存储单元 412 中还记录指向存储单元 411 的一个或多个指针。

图 5 是根据本发明实施例的主机执行写入操作的流程图。当用户应用程序或者其他程序请求执行将数据写入到存储设备的操作时，应用程序或其他程序会发送写请求。图 4 中的存储设备驱动程序 405 接收该写请求 (步骤 501)，该写请求中包括应用程序或其他程序所提供的要写入的数据已经用于存储设备的地址，用于存储设备的地址可以是文件路径及偏移值，并进一步被转换为用于存储设备的逻辑地址。在一个例子中，该逻辑地址是用于存储设备上的闪存芯片的逻辑地址，该写请求要将数据基于该逻辑地址写入到闪存芯片中。在步骤 502，存储设备驱动程序 405 为该写请求分配空闲的缓冲存储器。具体地，遍历缓冲区控制块 406，找到其中为空闲状态的存储单元，例如，存储单元 411。存储单元 411 为空闲状态，表示在存储设备 102 的缓冲存储器 106 中的对应存储单元为空闲状态，可以接收写入的数据。

在步骤 503 中，存储设备驱动程序 405 向存储设备 102 发送写命令，在写命令中包括要写入的数据以及用于存储设备 102 的地址，在该写命令中还包括同存储单元 411 相对应的用于缓冲存储器 106 的地址。所属领域技术人员将意识到，有多种方式获得存储单元 411 与缓冲存储器 106 中的对应存储单元的对应关系。例如，缓冲区控制块 406 中有 n 个存储单元 (411、412... ..41n)，而缓冲存储器 106 中也包括 n 个存储单元，存储单元 411 对应于缓冲存储器 106 中的第一个存储单元，而存储单元 412 对应于缓冲存储器 106 中的第二个存储单元，以及类似地，存储单元 41n 对应于缓冲存储器 106 中的第 n 个存储单元，使得基于存储单元 411 在缓冲区控制块 406 中的位置可计算出缓冲存储器 106 中的对应存储单元的地址。依然作为一个例子，还可以在存储单元 (411、412... ..412) 中存储缓冲存储器 106 中的对应存储单元的地址。在依然另一个例子中，在写命令中携带一个序号，该序号既指示存储单元 411 在缓冲区控制块 406 中的位置，又指示在缓冲存储器 106 中的对应存储单元的位置。

在步骤 504，接收到来自存储设备 102 的消息。在一个例子中，该消息是中断请求，该中断请求指示在步骤 503 中发送的写命令已经执行完毕。如前面所述的，在一个例子中，存储设备 102 中的控制电路 104 在将写命令中的数据写入到缓冲存储器 106 (特别地，写入到缓冲存储器 106 中与缓冲区控制块 406 的存储单元 411 相对应的存储单元) 之后，存储设备即发送中断，指示该写命令执行完成。在一个

例子中，控制电路 104 将数据写入到闪存芯片 105 之后，存储设备 102 向主机 101 发送中断。在一个例子中，该中断请求中还包括指示步骤 503 中发送的写命令相关的缓冲区控制块 406 的存储单元的信息。该信息可以是一个或多个存储单元（411、412... ..41n）的地址，一个或多个存储单元（411、412... ..41n）的序号。

在步骤 505，响应于在步骤 504 中接收到的该中断请求，并基于该中断请求中所指示的与该写命令相关的缓冲区控制块 406 的存储单元的信息，释放与该写命令相关的缓冲区控制块 406 的存储单元（411、412... ..41n）。释放存储单元（411、412... ..41n）具体可以是在存储单元（411、412... ..41n）中设置缓冲存储器 106 中的对应存储单元是空闲状态。

在一个例子中，步骤 503 中的写命令涉及缓冲区控制块 406 的两个存储单元 411 与 412，并且，存储单元 411 与 412 中分别记录了指向彼此的指针，以表示这两个存储单元 411、412 关联于同一个写命令。在步骤 504 中接收到的中断请求中，既可以指示存储单元 411 也可以指示存储单元 412。在步骤 505 中，基于指示存储单元 411 与 412 之一的指针，可获得两个存储单元 411 与 412，并将其释放。类似地，所属领域技术人员可意识到以此方式还可以在写命令中关联三个或更多的存储单元（411、412... ..41n）。

图 6 是根据本发明实施例的第二写入命令的示意图。第二写入命令指示存储设备 102 以 DMA 方式从主机 101 获得数据并写入到闪存芯片 105 中。第二写入命令可以是 DMA 描述符 600。DMA 描述符 600 包括 DMA 命令 610 以及一个或多个 DMA 数据（620、630）。DMA 命令 610 包括字段 611，用于指示 DMA 方式，即该 DMA 描述符 600 指示的操作，其可以为闪存读、写、擦除或者其他操作。字段 612 指示存储设备的逻辑地址。字段 613 指示该 DMA 描述符 600 的长度，即该 DMA 描述符 600 所包括的 DMA 数据（620、630）的个数，其可以为 1 个或多个。字段 614 指示缓冲存储器 106 的地址。DMA 数据 620、630 分别包括字段 621、631，用于指示 DMA 传输中的主机地址。DMA 数据 620、630 还分别包括字段 622、632，用于指示缓冲存储器 106 的地址。

DMA 命令 610 中的字段 612 的存储设备的逻辑地址可用于 DMA 数据 620、630。在 DMA 描述符 600 中仅包括一个 DMA 数据 620 的情况下，存储设备 102 根据字段 621 指示的主机地址和字段 622 指示的缓冲存储器地址，在主机 101 和存储设备 102 之间发起 DMA 传输，并最终将接收到的数据存储到由字段 612 所指示的闪存芯片 105 中。在 DMA 描述符 600 包括 DMA 数据 620 与 630 的情况下，存储设备 102 将根据 DMA 数据 620 而执行 DMA 传输所得的数据，最终存储在由字段 612 所指示的闪存芯片 105 中，而存储设备 102 还将根据 DMA 数据 630 而执行 DMA 传输所得的数据，最终存储在由字段 612 加上一预定偏移值所指示的闪存芯片 105 中。换句话说，DMA 描述符 600 可以指示在主机 101 和存储设备 102 之间的多次 DMA 传输，每次 DMA 传输同 DMA 数据 620、630 中的一个相对应，每次 DMA 传输中传输相同数量的数据（例如 4K 字节），并且 DMA 数据 620 与 DMA 数据 630 所对应的 DMA 传输的存储设备的逻辑地址是连续的（例如，相距预定的偏移值，该偏移值可以与 DMA 传输的数据量相对应）。这样，可以在 DMA 描述符 600 中仅携带一个存储设备逻辑地址（字段 612）。而 DMA 数据 620 与 DMA 数据

630 所对应的 DMA 传输的两个 DMA 主机地址 (字段 621、631) 可以是不连续的, 这样可以支持分散-收集 (Scatter-Gather) 方式的 DMA 传输。

在对应于 DMA 数据 620 的 DMA 传输中, 将来自字段 621 所指示的 DMA 主机地址的数据, 写入到字段 622 所指示的缓冲存储器 106 中, 继而再写入到闪存芯片 105 中。在对应于 DMA 数据 630 的 DMA 传输中, 将来自字段 631 所指示的 DMA 主机地址的数据, 写入到字段 632 所指示的缓冲存储器 106 中, 继而再写入到闪存芯片 105 中。

字段 614 是可选的。在字段 614 所对应的缓冲存储器地址处, 作为一个例子, 可保存字段 613 所指示的 DMA 描述符 600 的长度。从而可以记录 DMA 多个数据 620、630 所对应的 DMA 传输有多少已经得到执行, 或者有多少尚未被执行。对于 DMA 描述符 600, 当其全部 DMA 数据 620、630 所对应的 DMA 传输均已执行完毕后, 例如, 所对应的数据均写入到闪存芯片 105 中后, 存储设备 102 向主机 101 发送中断, 以指示对 DMA 描述符 600 的执行完成。这样, 对于 DMA 描述符 600, 虽然其对应于 2 次 DMA 传输过程, 但仅向主机 101 发送一次中断。减少中断请求次数, 将有助于降低主机 101 的工作负荷。

图 7A、7B 是根据本发明实施例的存储设备执行第二写入命令的流程图。第二写入命令可以是如图 6 所示的描述符 600。参看图 7A, 在步骤 701, 存储设备 102 接收 DMA 描述符 600。在 DMA 描述符 600 中包括用于主机的地址 (例如, DMA 主机地址 621、631), 用于闪存芯片 105 的地址 (例如, 存储设备逻辑地址 612) 以及用于缓冲存储器 106 的地址 (例如, 缓冲存储器地址 622、632)。虽然在图 6 中的 DMA 描述符 600 包括 DMA 命令 610、DMA 数据 620、630, 但这仅是为了清楚表达的目的。DMA 命令 610、DMA 数据 620、630 也可以组合在一起。存储设备 102 从 DMA 描述符 600 中提取出用于主机的地址, 用于闪存芯片 105 的地址以及用于缓冲存储器 106 的地址。

在步骤 702, 存储设备 102 基于用于主机的地址和用于缓冲存储器的地址, 以 DMA 传输方式从主机 101 将数据写入到缓冲存储器 106。在步骤 703, 基于用于闪存存储器的地址和用于缓冲存储器, 将在步骤 702 中写入到缓冲存储器中的数据, 写入到闪存芯片 105。

在一个例子中, 在步骤 702, 将数据写入到缓冲存储器 106 之后, 向主机 101 发送中断, 以指示对 DMA 传输的执行完成。如果 DMA 描述符 600 中仅包括这一次 DMA 传输 (例如, DMA 描述符 600 仅包括 DMA 命令 610 与 DMA 数据 620), 则该中断也表示对 DMA 描述符 600 的执行完成。在一个例子中, 在步骤 703, 将数据写入到闪存芯片 105 之后, 向主机 101 发送中断, 以指示对 DMA 传输的执行完成。

参看图 7B, 其示出了对包含多个 DMA 数据 (620、630) 的 DMA 描述符 600 的更详细的处理过程。在步骤 711, 存储设备 102 接收 DMA 描述符 600, DMA 描述符 600 包括 DMA 命令 610、DMA 数据 620 以及 DMA 数据 630。

在步骤 712, 基于 DMA 描述符 600, 将 DMA 数据 620 变换为第一 DMA 微指令, 将 DMA 数据 630 变换为第二 DMA 微指令。第一 DMA 微指令中包括 DMA 主机地址 621 和缓冲存储器地址 622。基于第一 DMA 微指令, 还可以获得与其对应

的存储设备的逻辑地址 612、DMA 操作类型以及 DMA 描述符长度 613。该存储设备的逻辑地址 612 和 DMA 操作类型可以是第一 DMA 微指令的一部分，也可以是存储在缓冲存储器 106 中，并通过在第一 DMA 微指令中的索引来访问，还可以通过将第一 DMA 微指令放置在特定的操作队列（读、写、擦除、其他）中，以标识第一 DMA 微指令的操作类型。

作为依然另一个例子，对于第一 DMA 微指令，根据缓冲存储器地址 622 与一预定的偏移值得到存储在缓冲存储器 106 中的指针，基于该指针，获得存储设备的逻辑地址 612 和/或 DMA 描述符长度 613。通过类似的方式，对于第二 DMA 微指令，根据缓冲存储器地址 632 与一预定偏移值得到存储在缓冲存储器 106 中的指针，基于该指针，获得存储设备的逻辑地址 612 和/或 DMA 描述符长度 613。

在优选的实施例中，基于缓冲存储器地址 614，将 DMA 描述符长度 613 存储在缓冲存储器 106 中，其中 DMA 描述符长度等同于 DMA 描述符 600 中的 DMA 数据的个数（或者 DMA 描述符 600 中的 DMA 命令与 DMA 数据的总计个数，从中可以得到 DMA 数据的个数），并通过在第一 DMA 微指令中的索引来访问 DMA 描述符长度 613。这样，第一 DMA 微指令与第二 DMA 微指令的执行顺序变得不重要。每执行一个 DMA 微指令，将缓冲存储器 106 中的 DMA 描述符长度递减（例如，减 1 或减去单位长度），当缓冲存储器 106 中的 DMA 描述符长度变为 0 时，表示对该 DMA 描述符 600 的所有 DMA 操作全都执行完成。

在步骤 713，对于第一 DMA 微指令，基于 DMA 主机地址 621 和缓冲存储器地址 622，以 DMA 传输方式将数据写入到相应缓冲存储器 106 中。

在步骤 714，对于第二 DMA 微指令，基于 DMA 主机地址 631 和缓冲存储器地址 632，以 DMA 传输方式将相应数据写入到缓冲存储器 106 中。

在步骤 715，向主机发送中断，以指示对 DMA 描述符 600 的操作完成。

在一个例子中，步骤 713，还包括将同第一 DMA 微指令相对应的写入到缓冲存储器 106 中的数据，基于存储设备逻辑地址 612，写入到闪存芯片 105 中。步骤 714 还包括将同第二 DMA 微指令相对应的写入到缓冲存储器 106 中的数据，基于存储设备逻辑地址 612 加上预定偏移值的和，写入到闪存芯片 105 中。

在优选的实施例中，在步骤 713、714 中，将相应数据写入到缓冲存储器 106 中之后，还基于在第一、第二微指令中的索引，访问存储在缓冲存储器中的 DMA 描述符长度 613，并将 DMA 描述符长度 613 递减（例如，减 1 或减去单位长度）。这样，当该 DMA 描述符长度 613 变为 0 时，意味着对 DMA 描述符 600 的操作完成。以此方式，可以在存储设备 102 中同时处理多个 DMA 描述符 600，并且第一 DMA 微指令与第二 DMA 微指令的执行顺序也是不重要的。还可以采用其他方式来识别 DMA 描述符 600 中的多个 DMA 数据均被执行。例如，在缓冲存储器或寄存器中为每个 DMA 描述符 600 的每个 DMA 数据提供标志，每当一个 DMA 数据（DMA 微指令）被执行后，将相应的标志置位。还可以顺序执行 DMA 描述符 600 中的每个 DMA 数据（DMA 微指令），当最后一个 DMA 数据（DMA 微指令）被执行后，意味着对该 DMA 描述符 600 的执行完成。

在依然另一个实施例中，当 DMA 描述符 600 中的每一个 DMA 微指令被执行后，向主机 101 发送中断，并由主机驱动程序来分析对 DMA 描述符 600 的执行是

否已完成。分析方法同上面所描述的存储设备 102 中识别 DMA 描述符 600 中的多个 DMA 数据是否均被执行的过程相类似。

图 7C 中展示了用于实施图 7A、7B 中的执行第二写入命令的存储设备的硬件方框图。图 7C 中，主机 101 包括 PCIE 控制器 721 和主机存储器 722。主机存储器 722 可以是随机访问存储器 (RAM)，PCIE 控制器 721 用于同存储设备经由 PCIE 总线进行通信。存储设备 102 中包括 PCIE 接口 731、DMA 指令分析器 732、微指令 FIFO (先进先出缓冲器) 733、DMA 写操作控制器 734、DMA 写接口 735、逻辑地址到物理地址转换电路 736、Flash 接口控制器 737、闪存芯片 105 以及缓冲存储器 106。

PCIE 接口 731 接收主机 101 通过 PCIE 控制器 721 发送的 DMA 描述符 600。主机 101 与存储设备 102 之间的连接不限于 PCIE 方式，还可以通过 SATA、IDE、USB、PCIE、SCSI、以太网、光纤通道等连接主机 101 与存储设备 102。DMA 指令分析器 732 将 PCIE 接口 731 接收到的 DMA 描述符 600 变换为 DMA 微指令。对于如图 6 所示的 DMA 描述符 600，其中包括 DMA 数据 620 和 DMA 数据 630，则 DMA 指令分析器将其变换为对应于 DMA 数据 620 的第一 DMA 微指令和对应于第二 DMA 数据 630 的第二 DMA 微指令。第一 DMA 微指令与第二 DMA 微指令的结构已在上文中详细介绍。DMA 指令分析器 732 还从 DMA 描述符 600 的 DMA 命令 610 中提取出 DMA 描述符长度，并对其加以保存，可以保存在缓冲存储器 106、一个寄存器或者类似物中。DMA 描述符长度指示了 DMA 描述符 600 所包括的 DMA 数据的数量，也指示了从该 DMA 描述符 600 所得到的 DMA 微指令的个数。DMA 指令分析器 732 将第一 DMA 微指令与第二 DMA 微指令存储在微指令 FIFO 733 中。

微指令 FIFO 733 能够缓存 DMA 微指令，并按照先进先出的方式向 DMA 写操作控制器 734 提供 DMA 微指令。虽然这里仅以写操作为例，描述了将与 DMA 写操作对应的 DMA 微指令缓存在微指令 FIFO733 中，所属领域技术人员将意识到可以将与 DMA 读操作对应的 DMA 微指令同与 DMA 写操作对应的 DMA 微指令混合缓存在微指令 FIFO733。还可以将微指令 FIFO733 配置为两部分或多个部分，其中一部分专用于存储与 DMA 读操作对应的 DMA 微指令，而将另一部分专用于存储与 DMA 写操作对应的 DMA 微指令。

对于与 DMA 写操作对应的 DMA 微指令，例如，前面所提到的第一 DMA 微指令与第二 DMA 微指令，DMA 写操作控制器 734 基于这些 DMA 微指令来执行 DMA 写操作。如同前面所提到的，第一 DMA 微指令中包括，DMA 主机地址 621 和缓冲存储器地址 622。DMA 写操作控制器 734 利用 DMA 主机地址 621 和缓冲存储器地址 622，通过 DMA 写接口 735 在主机 101 和存储设备 102 之间发起 DMA 写操作，将存储在 DMA 主机地址 621 处的数据，传输到缓冲存储器地址 622 所指示的位置，所传输的数据可以具有预定的长度 (例如 4K 字节)。对于第二 DMA 微指令，DMA 写操作控制器 734 执行类似的操作，将存储在 DMA 主机地址 631 处的数据，传输到缓冲存储器地址 632 所指示的位置。

从第一 DMA 微指令和第二 DMA 微指令可获得用于各自的存储设备的逻辑地址。在逻辑地址到物理地址转换电路 736，为每条 DMA 微指令的存储设备的逻辑

地址转换为用于闪存芯片 105 的物理地址。对于每一条 DMA 微指令，Flash 接口控制器 737 将写入到缓冲存储器 106 的数据，基于逻辑地址到物理地址转换电路 736 所提供的物理地址，写入到闪存芯片 105 中。

Flash 接口控制器 737 还基于在第一、第二微指令中的索引，访问所保存的从 DMA 描述符 600 的 DMA 命令 610 中提取出 DMA 描述符长度，并将 DMA 描述符递减（例如，减 1 或减去单位长度）。这样，当该 DMA 描述符长度变为 0 时，意味着对 DMA 描述符 600 的操作完成。继而，可向主机发送中断，以指示对 DMA 描述符 600 的操作完成。在一个例子中，DMA 写操作控制器也访问所保存的 DMA 描述符长度，并确定是否已将与 DMA 描述符 600 相对应的所有数据均写入到缓冲存储器 106 中，并向主机发送指示所有数据已写入到缓冲存储器 106 的中断。

图 8 是根据本发明实施例的主机的软件方框图。图 8 是示出了主机 800 的软件方框图，其同图 4 中展示的主机的软件方框图相类似。不同之处在于，图 8 中的缓冲区控制块 406 中还包括 IO 请求链表 801。IO 请求链表 801 是利用缓冲区控制块 406 中的存储单元（411、412... 41n）组成的链表。IO 请求链表 801 可以是单向链表、双向链表或循环链表。在生成如图 6 所示的 DMA 描述符 600 时，对于一个 DMA 描述符 600，创建一个与之相对应的 IO 请求链表 801，其中包括分别与 DMA 命令 610、DMA 数据 620、DMA 数据 630 分别相对应的存储单元（811、812、813）。需要指出的是，存储单元（811、812、813）是缓冲区控制块 406 的存储单元（411、412... 41n）中的三个存储单元，并通过设置相应的指针，形成 IO 请求链表 801。在图 8 中，将存储单元（811、812、813）与存储单元（411、412... 41n）分开展示，仅仅是为了清楚地描述的需要。

图 9A 是根据本发明实施例的主机创建并执行第二写入命令的流程图。在一个实施例中，步骤 901，由主机的存储设备驱动程序 405 接收 IO 请求。该 IO 请求指示将分散在主机存储器的不同物理地址的多个数据块写入到存储设备 102 中，为此将在主机与存储设备之间执行分散-收集 DMA 操作。下面将以举例的方式描述主机向存储器写入数据的操作过程。

在步骤 902，结合图 8，从缓冲区控制块 406 中取出一个空闲状态的存储单元，例如存储单元 411。

在步骤 904，根据 IO 请求的内容，创建 DMA 描述符 600 的 DMA 命令 610，填充 DMA 命令 610 中的 DMA 方式字段 611（在该例子中，是写操作）、存储设备逻辑地址字段 612（从 IO 请求中可获得该信息）、DMA 描述符长度字段 613（从 IO 请求中可获得该信息）以及缓冲存储器地址字段 614（与步骤 902 中所分配的存储单元 411 相对应）。继而将所创建的 DMA 命令 610 发送给存储设备 102。并将存储单元 411 作为用于该 IO 请求的 IO 请求链表 801 的起始节点（例如存储单元 811）。创建 IO 请求链表 801，用于在存储设备 102 执行完 DMA 描述符 600 之后，将所占用的存储单元归还给缓冲区控制块 406，并就该 IO 请求的执行完成通知应用软件或其他上层软件。为此目的，在一个例子中，还在存储单元 411 中存储对应于该 IO 请求的指针。基于 DMA 描述符长度，还可获得 DMA 数据部分的剩余长度。在生成第一个 DMA 数据之前，DMA 数据部分的剩余长度是 DMA 描述符 600 中的 DMA 数据（620、630）的个数，作为一个例子，其为 DMA 描述符长度减 1。

在步骤 906, 从缓冲区控制块 406 中取出一个空闲状态的存储单元, 例如, 存储单元 412。根据 IO 请求的内容, 创建 DMA 描述符 600 的 DMA 数据 620, 填充 DMA 数据 620 中的 DMA 主机地址字段 621 (从 IO 请求中可获得该信息) 以及缓冲存储器地址 622 (与所分配的存储单元 412 相对应, 例如, 存储单元 412 在缓冲区控制块 406 中的偏移值或序号)。继而将所创建的 DMA 数据 620 发送给存储设备 102。并将存储单元 412 作为用于该 IO 请求的 IO 请求链表 801 的节点 (例如存储单元 812)。

在步骤 908, 将 DMA 数据部分剩余长度递减, 得到 DMA 描述符 600 中尚未发送给存储设备的 DMA 数据的个数。

在步骤 909, 如果 DMA 数据部分的剩余长度为 0, 则表示 DMA 描述符 600 的生成已经完成, 进而在步骤 910 存储设备驱动程序 405 将等待存储设备 102 返回的表示 DMA 描述符 600 的处理已经完成的中断, 并依据该中断找到与之对应的 IO 请求链表 801, 以及将 IO 请求链表 801 中的存储单元 (811、812) 释放。换句话说, 将由 IO 请求链表 801 中的存储单元 (811、812) 的状态设置为空闲, 从而使得通过缓冲区控制块 406 可以获知存储单元 411、412 的状态为空闲。在一个例子中, 在 DMA 描述符 600 指示读操作, 且主机 101 的 CPU 包括高速缓冲存储器的情况下, 还通知同 DMA 描述符 600 的 DMA 主机地址 (621、631) 相关联的 CPU 高速缓冲存储器执行一致性处理, 以反映出 DMA 主机地址 (621、631) 处的数据可能因读 DMA 读操作而发生变化。在一个例子中, 存储设备 102 返回的中断中包括指示 IO 请求链表 801 中的多个存储单元 (811、812) 之一 (或者缓冲存储器地址 622、632 之一) 的内容, 依据该内容, 通过 IO 请求链表 801 将存储单元 (811、812) 释放。

在步骤 909, 如果 DMA 数据部分的剩余长度大于 0, 则表示 DMA 描述符 600 的生成尚未完成, 还需要为 IO 请求生成一个或多个 DMA 数据, 那么处理将返回到步骤 906 并重复执行步骤 906、908 和 909。

图 9B 是根据本发明另一实施例的主机创建并执行第二写入命令的流程图。在该实施例中, 将缓冲区控制块 406 中的空闲存储单元进一步组织为空闲存储单元池, 以有助于 DMA 描述符 600 的创建过程。通过将缓冲区控制块 406 中状态为空闲的存储单元 (411、412... 41n) 组织成链表来形成空闲存储单元池。当需要从缓冲区控制块 406 中获得空闲存储单元时, 可以从空闲存储单元池中取出存储单元, 从而省去了在缓冲区控制块 406 中查找空闲存储单元的开销。

在一个实施例中, 步骤 921, 由主机的存储设备驱动程序 405 接收 IO 请求。同图 9A 相类似, 该 IO 请求指示将分散在主机存储器的不同物理地址的多个数据块写入到存储设备 102 中, 为此将在主机与存储设备之间执行分散-收集 DMA 操作。

在步骤 922, 根据 IO 请求的内容, 计算出用于与相同该 IO 请求相对应的 DMA 描述符的长度 (例如, DMA 命令以及 DMA 数据的个数)。注意到在图 9A 公开的实施例中, 是在步骤 904 创建 DMA 命令 610 的过程中获得 DMA 描述符长度, 所属领域技术人员将意识到各个步骤并非必然以本实施例中所公开的顺序执行。

在步骤 923, 判断空闲存储单元池中是否为空。如果空闲存储单元池非空, 即缓冲区控制块 406 中存在处于空闲状态的存储单元, 则进行到步骤 924, 并从空闲

存储单元池中取出一个空闲存储单元（例如存储单元 411）。如果空闲存储单元池为空，意味着缓冲区控制块 406 中没有空闲的存储单元。则在步骤 925，等待空闲存储单元池被更新，以出现空闲的存储单元。当 DMA 描述符的执行完成后，与之相关的存储单元会被释放，从而在空闲存储单元池中出现空闲的存储单元。后面对此会详细介绍。

在步骤 926，确定当前要生成用于 DMA 描述符 600 的 DMA 命令字段还是 DMA 数据字段。一般而言，DMA 描述符 600 包括一个 DMA 命令和一个或多个 DMA 数据。当要生成 DMA 命令时，处理转向步骤 927，并根据 IO 请求的内容，创建 DMA 描述符 600 的 DMA 命令 610，填充 DMA 命令 610 中的各个字段（611、612、613、614）。在一个例子中，还在存储单元 411 中存储对应于该 IO 请求的指针，以便在该 IO 请求的执行完成后，可识别该 IO 请求并通知应用软件或其他上层软件。当要生成 DMA 数据时，处理转向步骤 928，并根据 IO 请求的内容，创建 DMA 描述符 600 的 DMA 数据 620，填充 DMA 数据 620 的各个字段（621、622）。

接下来，在步骤 929，将所生成的 DMA 命令或 DMA 数据发送给存储设备 102。并在步骤 930，将在步骤 924 中获得的存储单元 411 设置在 IO 请求链表 801 中。作为一个例子，第一个进入到 IO 请求链表 801 的存储单元，将作为 IO 请求链表 801 的头节点，但是，也将意识到当 IO 请求链表 801 被组织为环形链表时，其中并不存在“头节点”。在步骤 930，还将 DMA 描述符长度递减。

在步骤 931，如果 DMA 描述符长度为 0，意味着 DMA 描述符 600 的生成已经完成，进而在步骤 932 存储设备驱动程序 405 将等待存储设备 102 返回的表示 DMA 描述符 600 的处理已经完成的中断，并依据该中断找到与之对应的 IO 请求链表 801，以及将 IO 请求链表 801 中的存储单元（811、812）释放。换句话说，将由 IO 请求链表 801 中的存储单元（811、812）的状态设置为空闲，从而使得通过缓冲区控制块 406 可以获知存储单元 411、412 的状态为空闲，并将存储单元 411、412 放入空闲存储单元池中。在一个例子中，存储设备 102 返回的中断中包括指示 IO 请求链表 801 中的多个存储单元（811、812）之一的内容，依据该内容，通过 IO 请求链表 801 将存储单元（811、812）释放。

在步骤 931，如果 DMA 描述符长度大于 0，则表示 DMA 描述符 600 的生成尚未完成，还需要为 IO 请求生成一个或多个 DMA 数据，那么处理将返回到步骤 923 并重复执行步骤 923-931。

上面结合图 9A、9B 描述了 DMA 描述符 600 的生成过程。DMA 描述符 600 用于在分散-收集 DMA 中描述要执行的多个 DMA 操作，该多个 DMA 操作的数据来源于存储在连续或不联系的存储空间中。所属领域技术人员将容易意识到，DMA 描述符 600 的生成方式包括但不限于上面图 9A、9B 中描述的具体方式。

图 10A 是根据本发明的实施例的在存储设备的缓冲存储器中创建链表的流程图。在如图 7A-7C 所公开的存储设备处理 DMA 描述符 600 的过程中，将 DMA 描述符 600 转换为一个或多个微指令。在进一步的实施例中，为了有效处理一个或多个微指令之间的关联关系（例如，这些微指令均同 DMA 描述符 600 相关联），存储设备 102 响应于主机 101 所传输的 DMA 描述符 600，还在缓冲存储器 106 中建立链表，该链表将对应于同一 DMA 描述符 600 的多个微指令关联起来。

如图 10A 所示, 在步骤 1002, 主机 101 向存储设备 102 发送 DMA 描述符 600。DMA 描述符 600 包括 DMA 命令 610 与 DMA 数据 620、630。前面已经结合图 9A 与图 9B 而描述了主机 101 向存储设备 102 发送 DMA 描述符 600 的过程的例子。还应当意识到, 在存储设备的缓冲存储器中创建链表, 将有助于存储设备对 IO 操作的执行, 特别是对多个 IO 操作的并发/乱序执行, 多个 IO 操作可通过访问各自的链表而关联在一起。这样不具备关联关系的 IO 操作可以在存储设备中并发执行。因而, 还可以响应除 DMA 命令之外的其他类型的 IO 命令或其他命令, 以在存储设备中创建链表。

在步骤 1004, 判断所接收到的是 DMA 命令 610 还是 DMA 数据 620、630。

如果接收到 DMA 命令 610, 在步骤 1006, 从其中的缓冲存储器地址字段 610 中提取出用于该 DMA 命令 610 的缓冲存储器地址, 并基于该缓冲存储器地址, 为该 DMA 命令 610 在缓冲存储器 106 中分配存储空间。接下来, 在步骤 1008, 保存为该 DMA 命令 610 所分配的缓冲存储器地址, 用来在为 DMA 数据 620、630 分配缓冲存储器地址时使用。

如果在步骤 1004 判断出所接收到的是 DMA 数据 620, 则在步骤 1010, 从 DMA 数据 620 的缓冲存储器地址字段 622 中提取出用于该 DMA 数据 620 的缓冲存储器地址, 并基于该缓冲存储器地址, 为该 DMA 数据 620 在缓冲存储器 106 中分配存储空间。并在步骤 1012 中, 在该 DMA 数据 620 所分配的缓冲存储器的存储空间中, 存储在步骤 1008 中保存的 DMA 命令 610 的缓冲存储器地址。这样, 在缓冲存储器 106 中, 为 DMA 命令 610 和 DMA 数据 620 所分配的存储空间形成了链表, 其中为 DMA 命令 610 所分配的存储空间是链表的头节点, 为 DMA 数据 620 所分配的存储空间连接到链表的头节点。

在 DMA 描述符 600 还包括 DMA 数据 630 的情况下, 通过步骤 1010 和步骤 1012, 基于 DMA 数据 630 中的缓冲存储器地址 632 为 DMA 数据 630 在缓冲存储器 106 中分配存储空间, 并在为 DMA 数据 630 所分配的缓冲存储器的存储空间中, 保存 DMA 命令 610 的缓冲存储器地址。所属领域技术人员将意识到, 也可以在为 DMA 数据 630 所分配的缓冲存储器 106 的存储空间中, 保存用于 DMA 数据 620 的缓冲存储器地址, 从而形成不同类型的链表。在其他例子中, 将缓冲存储器 106 中为 DMA 命令 610、DMA 数据 620、630 所分配的存储空间创建为循环链表或双向链表。

上面结合图 7A、7B、7C 已经描述了存储设备 102 基于 DMA 数据 (620、630) 生成 DMA 微指令, 并保存在微指令 FIFO 733 中。DMA 数据 (620、630) 生成 DMA 微指令的操作可以发生于步骤 1012 之后, 并在 DMA 微指令中携带为 DMA 数据 (620、630) 所分配的缓冲存储器地址。

图 10B 是根据本发明的一实施例的在存储设备的缓冲存储器中创建链表的流程图。同图 10A 所提供的实施例相比, 图 10B 的实施例中, 还将与 DMA 描述符的处理或执行相关的信息存储在所创建的链表中。图 10C 是根据本发明的一实施例的存储设备利用在缓冲存储器中创建的链表执行 DMA 描述符的流程图。图 11A-11F 展示出了与图 10B 与图 10C 相关联的缓冲存储器的多种状态。在图 11A-11F 中, 1100 指示缓冲存储器 106 中的存储空间。

具体地，在步骤 1020，主机 101 向存储设备 102 发送 DMA 描述符 600。在步骤 1022，判断所接收到的是 DMA 命令 610 还是 DMA 数据 620、630。

如果接收到 DMA 命令 610，在步骤 1024，从缓冲存储器地址字段 610 中提取出缓冲存储器地址，并基于该缓冲存储器地址，为该 DMA 命令 610 在缓冲存储器 106 中分配存储空间。参看图 11A，为 DMA 命令 610 分配存储空间 1101。以及还从 DMA 命令 610 中提取出 DMA 描述符长度字段 613，从 DMA 描述符长度 613 可以得到该 DMA 描述符 600 的 DMA 数据部分的长度（例如，DMA 描述符长度减 1）。接下来，在步骤 1026，保存为该 DMA 命令 610 所分配的缓冲存储器地址，用来在为 DMA 数据 620、630 分配缓冲地址时使用。并且，将 DMA 数据部分的长度记录在为该 DMA 命令 610 所分配的缓冲存储器中。参看图 11A，在存储空间 1101 中保存了 DMA 数据部分的长度（在这个例子中，DMA 数据部分的长度为 2）。

如果在步骤 1022 判断出所接收到的是 DMA 数据 620，则在步骤 1028，从 DMA 数据 620 的缓冲存储器地址字段 622 中提取出用于该 DMA 数据 620 的缓冲存储器地址，并基于该缓冲存储器地址，为该 DMA 数据 620 在缓冲存储器 106 中分配存储空间。参看图 11B，为 DMA 数据 620 分配存储空间 1112。并在步骤 1030 中，在为该 DMA 数据 620 所分配的缓冲存储器的存储空间 1112 中，存储在步骤 1026 中保存的 DMA 命令 610 的缓冲存储器地址。这样，在缓冲存储器 106 中，为 DMA 命令 610 和 DMA 数据 620 所分配的存储空间（1101 与 1112）形成了链表，其中为 DMA 命令 610 所分配的存储空间 1101 是链表的头节点，为 DMA 数据 620 所分配的存储空间 1112 连接到链表的头节点。还在存储空间 1112 中存储同 DMA 数据 620 相对应的 DMA 主机地址。

在 DMA 描述符 600 还包括 DMA 数据 630 的情况下，通过步骤 1028 和步骤 1030，基于 DMA 数据 630 中的缓冲存储器地址 632 为 DMA 数据 630 在缓冲存储器 106 中分配存储空间 1123（参看图 11C），并在存储空间 1123 中，保存 DMA 命令 610 的缓冲存储器地址。以及还在存储空间 1123 中存储同 DMA 数据 630 相对应的 DMA 主机地址。

因而，在缓冲存储器 106 中形成了同 DMA 描述符 600 相对应的链表，其中存储空间 1101 是链表的头节点，存储空间 1112 和 1123 是链表的节点，并指向该链表的头节点。所属领域技术人员将意识到，也可以在为 DMA 数据 630 所分配的缓冲存储器 106 的存储空间 1123 中，保存用于 DMA 数据 620 的缓冲存储器地址，从而形成不同类型的链表。在其他例子中，将缓冲存储器 106 中为 DMA 命令 610、DMA 数据 620、630 所分配的存储空间创建为循环链表或双向链表。

图 10C 是根据本发明的一实施例的存储设备利用在缓冲存储器中创建的链表执行 DMA 描述符的流程图。上面结合图 7A、7B、7C 已经描述了存储设备 102 基于 DMA 数据（620、630）生成 DMA 微指令，并保存在微指令 FIFO 733 中。在存储设备 102 对 DMA 微指令的执行中，在一个例子中，利用在缓冲存储器 106 中的链表。在 DMA 微指令中，包括缓冲存储器地址，通过该缓冲存储器地址，可以获得为同该 DMA 微指令相对应的 DMA 数据所分配的缓冲存储器 106 中的存储

空间，进而可以获得与该 DMA 数据所对应的 DMA 主机地址以及与该 DMA 数据所对应的 DMA 描述符中的 DMA 数据部分长度或 DMA 数据的个数。

在下面的描述中，将对应于 DMA 数据 620 的 DMA 微指令用第一 DMA 微指令指示，将对应于 DMA 数据 630 的 DMA 微指令用第二 DMA 微指令指示。

在步骤 1040，从微指令 FIFO733 中获得将第一 DMA 微指令。

在步骤 1042，第一 DMA 微指令中包括为 DMA 数据 620 所分配的缓冲存储器 106 的存储空间 1112 的地址，并从存储空间 1112 中获得 DMA 主机地址。该 DMA 主机地址是由 DMA 数据 620 中的 DMA 主机地址字段 621 所提供的。基于 DMA 主机地址，在主机 101 和存储设备 102 之间进行 DMA 传输，将主机 101 的该 DMA 主机地址处的预定长度（例如，4KB）的数据，以 DMA 方式传输到存储设备 102 的缓冲存储器中。对于第二 DMA 微指令，执行类似的操作，将主机 101 的由 DMA 数据 630 的 DMA 主机地址字段 632 所提供的 DMA 主机地址处的数据，以 DMA 方式传输到存储设备 102 的缓冲存储器的为 DMA 数据 632 所分配的存储空间 1123 处。在图 11D 中，示出了执行完第一 DMA 微指令与第二 DMA 微指令后，存储了以 DMA 方式传输的数据的存储空间 1112 和存储空间 1123。

在步骤 1044，继续对第一 DMA 微指令加以执行。通过第一 DMA 微指令中的存储空间 1112 的地址，从存储空间中取出预定长度的数据，该数据是在步骤 1042 中，通过 DMA 操作从主机 101 传输到缓冲存储器 106 的存储空间 1112 的。并通过闪存接口控制器（例如，图 7C 中的 Flash 接口控制器 737）将该数据基于第一 DMA 微指令中所包括的用于闪存存储器的地址，写入到闪存芯片 105。该用于闪存存储器的地址是通过 DMA 命令 610 中的存储设备逻辑地址字段 612 所得到的。对第二 DMA 微指令以类似的方式加以执行。通过第二 DMA 微指令中包括的存储空间 1123 的地址，从存储空间中取出预定长度的数据，并将该数据通过闪存接口控制器，基于第二 DMA 微指令中所包括的用于闪存存储器的地址，写入到闪存芯片 105。第二 DMA 微指令中所包括的用于闪存存储器的地址，是通过 DMA 命令 610 中的存储设备逻辑地址字段 612 加上预定值（例如对应于 DMA 传输的数据的长度，在该例子中，是 4KB）所得到的。在一个例子中，将 DMA 命令 610 中的存储设备逻辑地址转换为存储设备的物理地址，并基于该物理地址将数据写入到闪存芯片 105 中。从存储设备的逻辑地址到物理地址的映射过程，是所属领域技术人员所熟知的。

在步骤 1046，继续对第一 DMA 微指令加以执行。通过第一 DMA 微指令中的存储空间 1112 的地址，获得为 DMA 命令 610 所分配的存储空间 1101 的地址，并在存储空间 1101 中获得 DMA 数据部分长度，以及将存储 1101 中存储的 DMA 数据部分长度递减（例如，减 1 或者减去单位长度）。参看图 11E，对于第一 DMA 微指令，将存储空间 1101 中的 DMA 数据部分长度递减后，其值由 2 变为 1。并且，存储空间 1112 中不再保存存储空间 1101 的地址，用于表明对 DMA 数据 620 的执行已经完成。在步骤 1048，由于 DMA 数据部分的长度不为 0，意味着对 DMA 描述符 600 的操作尚未完成，因为其还包含另一个 DMA 数据 630，此时，不进行进一步的处理。

当第二 DMA 微指令在步骤 1046 被执行时，通过第二 DMA 微指令中的存

存储空间 1123 的地址, 获得为 DMA 命令 610 所分配的存储空间 1101 的地址, 并在存储空间 1101 中获得 DMA 数据部分长度, 以及将存储空间 1101 中存储的 DMA 数据部分长度递减 (例如, 减 1 或者减去单位长度)。参看图 11F, 对于第二 DMA 微指令, 将存储空间 1101 中的 DMA 数据部分长度递减后, 其值由 1 变为 0。并且, 存储空间 1123 中不再保存存储空间 1101 的地址, 用于表明对 DMA 数据 630 的执行已经完成。

此时, 当第二 DMA 描述符在步骤 1048 被执行时, 由于 DMA 数据部分的长度为 0, 意味着对 DMA 描述符 600 的执行已经完成。接下来, 在步骤 1050, 向主机 101 发送中断, 以指示对 DMA 描述符 600 的执行已经完成。

再次参看图 11F, 存储空间 1112 与 1123 均不再保存存储空间 1101 的地址。存储空间 1101 中的 DMA 数据部分长度的值为 0。在此情况下, 意味着对 DMA 描述符 600 的执行已经完成, 对存储空间 1101、1112 与 1123 均不会再加上使用, 这些存储空间可以被释放以用于对其他 DMA 描述符的执行。在一个例子中, 由主机 101 控制对相应存储空间的释放和再利用, 在上面已结合图 9A 与图 9B 对 IO 请求链表 801 中的存储空间的释放。由于缓冲区控制块 406 中的存储单元 (411、412... 41n) 与缓冲存储器 106 中的存储空间相对应, 因而对 IO 请求链表 801 中的存储空间的释放, 意味着对缓冲存储器中的存储空间 1101、1112 与 1123 的释放。

上面结合图 10B、10C、11A-11F 而描述了在缓冲存储器 106 中存储对应于第一与第二微指令的 DMA 主机地址的方案, 从而使得 DMA 微指令中不必携带 DMA 主机地址而减少了对电路资源的占用, 并通过缓冲存储器 106 将对应于同一 DMA 描述符 600 的第一、第二 DMA 微指令关联在一起的方案。所属领域技术人员将意识到, 还可以将对应于第一与第二微指令的存储设备逻辑地址和/或 DMA 主机地址存储在缓冲存储器中, 从而进一步减少 DMA 微指令的长度及其对电路资源的占用。

图 12 是根据本发明的又一实施例的存储设备的硬件方框图。与图 7C 中公开的相似, 主机 101 包括 PCIE 控制器 721 和主机存储器 722。存储设备 102 中包括 PCIE 接口 731、DMA 指令分析器 732、微指令先进先出缓冲器 (FIFO) 733、DMA 写操作控制器 734、DMA 写接口 735、逻辑地址到物理地址转换电路 736 以及缓冲存储器 106。存储设备 102 中还包括 DMA 读写微指令判断电路 1210, DMA 读接口 1212, 闪存控制器 1221、1222、1223, 闪存接口 1231、1232、1233, 完成控制电路 1242、多路选择器 1241 以及多路共享器 1243。闪存接口 1231、1232、1233 耦合于闪存芯片 105。

PCIE 接口 731 接收主机 101 通过 PCIE 控制器 721 发送的 DMA 描述符 600。主机 101 与存储设备 102 之间的连接不限于 PCIE 方式。DMA 指令分析器 732 将 PCIE 接口 731 接收到的 DMA 描述符 600 变换为 DMA 微指令。对于如图 6 所示的 DMA 描述符 600, 则 DMA 指令分析器将其变换为对应于 DMA 数据 620 的第一 DMA 微指令和对应于 DMA 数据 630 的第二 DMA 微指令。在一个例子中, 第一与第二 DMA 微指令中分别包括指示该微指令类型 (读/写/擦除/其他) 的字段、指示与其对应的缓冲存储器 106 中的存储单元的地址的字段、指示存储设备的逻辑地址的字段。

参看图 10B 与图 11C, DMA 指令分析器 732 还针对 DMA 命令 610, 在缓冲存储器 106 中分配存储单元, 并在其中存储 DMA 数据部分的长度。DMA 指令分析器还针对 DMA 数据 620, 在缓冲存储器 106 中分配存储单元, 并在其中存储为 DMA 命令 610 所分配的存储单元的地址, 以及存储 DMA 数据 620 中的 DMA 主机地址。DMA 指令分析器还针对 DMA 数据 630, 在缓冲存储器 106 中分配存储单元, 并在其中存储为 DMA 命令 610 所分配的存储单元的地址, 以及存储 DMA 数据 630 中的 DMA 主机地址。

DMA 指令分析器 732 将第一 DMA 微指令与第二 DMA 微指令存储在微指令 FIFO 733 中。

微指令 FIFO 733 能够缓存 DMA 微指令, 并按照先进先出的方式向 DMA 读写微指令判断电路 1210 提供 DMA 微指令。

在 DMA 读写微指令判断电路 1210, 判断所获得的 DMA 微指令的类型。对于与 DMA 写操作对应的 DMA 微指令, 例如, 前面所提到的第一 DMA 微指令与第二 DMA 微指令, DMA 写操作控制器 734 基于这些 DMA 微指令来执行 DMA 写操作。DMA 写操作控制器 734 利用第一 DMA 微指令中的指示与其对应的缓冲存储器 106 中的存储单元的地址的字段, 从缓冲存储器 106 中获得 DMA 主机地址, 并通过 DMA 写接口 735 在主机 101 和存储设备 102 之间发起 DMA 写操作, 将存储在 DMA 主机地址中的数据, 传输到与第一 DMA 微指令相对应的缓冲存储器 106 的存储单元中, 所传输的数据可以具有预定的长度 (例如 4K 字节)。对于第二 DMA 微指令, DMA 写操作控制器 734 利用与其对应的缓冲存储器 106 中的存储单元的地址的字段, 从缓冲存储器 106 中获得 DMA 主机地址, 并在主机 101 和存储设备 102 之间发起 DMA 写操作, 将存储在 DMA 主机地址中的数据, 传输到与第二 DMA 微指令相对应的缓冲存储器 106 的存储单元中。

从第一 DMA 微指令和第二 DMA 微指令的指示存储设备的逻辑地址的字段获得用于各自的存储设备的逻辑地址。在逻辑地址到物理地址转换电路 736, 将每条 DMA 微指令的存储设备的逻辑地址转换为用于闪存芯片 105 的物理地址。对于每一条 DMA 微指令, Flash 控制器 1221、1222、1223 基于逻辑地址到物理地址转换电路 736 所提供的物理地址, 通过闪存接口 1231、1232、1233 将写入到缓冲存储器 106 的数据, 写入到闪存芯片 105 中, 其中, Flash 控制器 1221 同闪存接口 1231 相耦合, Flash 控制器 1222 同闪存接口 1232 相耦合, Flash 控制器 1223 同闪存接口 1233 相耦合。而闪存接口 1231、1232、1233 分别耦合到各自的闪存芯片。因而对于从 DMA 微指令中的存储设备的逻辑地址转换得到的用于闪存芯片 105 的物理地址, 该物理地址指示了特定的闪存芯片, 并且该闪存芯片与闪存接口 1231、1232、1233 的特定一个相耦合。因而, 基于该物理地址, 可以确定使用闪存接口 1231、1232、1233 中的哪一个将数据写入闪存芯片, 也可以确定使用 Flash 控制器 1221、1222、1223 中的哪一个。闪存接口 1231、1232、1233 还通过多路选择器 1241 与缓冲存储器相耦合。基于该物理地址, 多路选择器 1241 将数据从缓冲存储器 106 传送给闪存接口 1231、1232、1233 中的特定一个。

闪存接口 1231、1232、1233 将数据写入到闪存芯片 105 中之后, 完成控制电路 1242 还基于在第一、第二微指令中的指示与其对应的缓冲存储器 106 中的

存储单元的地址的字段，访问缓冲存储器 106，并进而访问缓冲存储器中为 DMA 命令 610 所分配的存储单元，从中获得 DMA 数据部分的长度，并将 DMA 数据部分的长度递减（例如，减 1 或减去单位长度）。这样，当为 DMA 命令 610 所分配的存储单元中的该 DMA 数据部分的长度变为 0 时，意味着对 DMA 描述符 600 的操作完成。继而，可向主机发送中断，以指示对 DMA 描述符 600 的操作完成。

虽然图 12 中以举例的方式展示了包括三个 Flash 控制器 1221、1222、1223 以及三个闪存接口 1231、1232、1233 的实施例，所属领域技术人员将意识到可以使用多种不同数量的 Flash 控制器与闪存接口，以同闪存芯片的数量相适应。

对于与 DMA 读操作对应的 DMA 微指令，DMA 读写微指令判断电路 1210 将其直接传送给逻辑地址到物理地址转换电路 736，并得到用于闪存芯片 105 的物理地址。Flash 控制器 1221、1222、1223 基于该物理地址，通过闪存接口 1231、1232、1233 将从闪存芯片 105 中读出数据。并基于 DMA 微指令中的指示与其对应的缓冲存储器 106 中的存储单元的地址的字段，从缓冲存储器 106 中获得与该 DMA 微指令相对应的 DMA 主机地址，以及经由 DMA 读接口 1212 在主机 101 与存储设备 102 之间发起 DMA 传输，将读出数据传输到主机 101 的主机 RAM 722 中由该 DMA 主机地址所指示的位置处。闪存接口 1231、1232、1233 经由多路共享器 1243 与 DMA 读接口 1212 相耦合，使得从闪存接口 1231、1232、1233 获得的数据均可以通过 DMA 读接口 1212 传输到主机 RAM 722。当闪存接口 1231、1232、1233 从闪存芯片 105 读出数据后，完成控制电路 1242 也基于在 DMA 微指令中的指示与其对应的缓冲存储器 106 中的存储单元的地址的字段，访问缓冲存储器 106，并进而获得 DMA 数据部分的长度，并将 DMA 数据部分的长度递减（例如，减 1 或减去单位长度）。这样，当该 DMA 数据部分的长度变为 0 时，意味着对该 DMA 描述符的操作完成。继而，可向主机发送中断，以指示对该 DMA 描述符的操作完成。

在图 12 所公开的存储设备中，可以支持对多个 DMA 描述符 600 的并发操作。对于多个 DMA 描述符 600 的每一个，通过在缓冲存储器 106 中创建的链表，将与一个 DMA 描述符 600 相对应的 DMA 数据关联在一起，使得对多个 DMA 微指令的操作顺序变得不重要。

上面已经详细描述了存储设备执行数据写入操作或与写操作相关的 DMA 描述符的执行。显然，存储设备执行读操作也可以从本发明公开中获益。例如，可以在读取命令中同时指定闪存芯片地址和缓冲存储器地址，并利用缓冲存储器作为读出数据的缓冲。也可以在与读操作相关的 DMA 描述符中描述缓冲存储器地址，在将数据从闪存芯片中读出后，可利用缓冲存储器作为读出数据的缓存。

已经为了示出和描述的目的而展现了对本发明的描述，并且不旨在以所公开的形式穷尽或限制本发明。对所属领域技术人员，许多调整 and 变化是显而易见的。

权 利 要 求 书

1. 一种在信息处理设备和存储设备之间进行 DMA 传输的方法, 所述存储设备包括缓冲存储器和闪存芯片, 所述方法包括:

接收第一 IO 请求;

为所述第一 IO 请求分配第一存储单元与第二存储单元;

向所述存储设备发送所述第一 DMA 描述符, 所述第一 DMA 描述符包括用于信息处理设备的地址、用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址与第二用于缓冲存储器的地址, 其中, 所述第一用于缓冲存储器的地址同所述第一存储单元相对应, 所述第二用于缓冲存储器的地址同所述第二存储单元相对应;

在所述存储设备和所述信息处理设备之间依据所述第一 DMA 描述符进行 DMA 传输; 接收来自所述存储设备的消息, 所述消息指示所述存储设备对第一 DMA 描述符已执行完成;

释放所述第一存储单元与所述第二存储单元。

2. 根据权利要求 1 所述的方法, 其中分配第一存储单元与第二存储单元进一步包括, 在所述第二存储单元中存储指向所述第一存储单元的地址。

3. 一种在信息处理设备和存储设备之间进行 DMA 传输的方法, 所述存储设备包括缓冲存储器和闪存芯片, 所述方法包括:

接收第一 IO 请求;

为所述第一 IO 请求分配第一存储单元与第二存储单元;

向所述存储设备发送所述第一 DMA 描述符, 所述第一 DMA 描述符包括第一 DMA 描述符命令和第一 DMA 描述符数据, 所述第一 DMA 描述符命令包括用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址, 所述第一 DMA 描述符数据包括用于信息处理设备的地址以及第二用于缓冲存储器的地址; 其中, 所述第一用于缓冲存储器的地址同所述第一存储单元相对应, 所述第二用于缓冲存储器的地址同所述第二存储单元相对应;

在所述存储设备和所述信息处理设备之间依据所述用于信息处理设备的地址和所述第二用于缓冲存储器的地址进行 DMA 传输;

接收来自所述存储设备的消息, 所述消息指示所述存储设备对第一 DMA 描述符已执行完成;

释放所述第一存储单元与所述第二存储单元。

4. 根据权利要求 1-3 之一所述的方法, 其中, 所述第一 IO 请求指示将所述用于信息处理设备的地址处的数据写入到用于所述存储设备的闪存芯片的地址处。
5. 根据权利要求 1-3 之一所述的方法, 还包括将空闲的存储单元链接在一起形成空闲存储单元池, 在分配第一存储单元与第二存储单元时, 从所述空闲存储单元池中取出空闲的存储单元, 在释放所述第一存储单元与所述第二存储单元时, 将所述第一存储单元与所述第二存储单元插入到所述空闲存储单元池。
6. 根据权利要求 1-5 之一所述的方法, 还包括将空闲的存储单元链接在一起形成空闲存储单元池, 在分配第一存储单元与第二存储单元时, 若所述空闲存储单元池为空, 则等待一个或多个存储单元被插入到所述空闲存储单元池。
7. 一种在信息处理设备和存储设备之间进行 DMA 传输的方法, 所述存储设备包括缓冲存储器和闪存芯片, 所述方法包括:

接收第一 IO 请求;

为所述第一 IO 请求分配第一存储单元、第二存储单元与第三存储单元;

向所述存储设备发送所述第一 DMA 描述符, 所述第一 DMA 描述符包括第一 DMA 描述符命令、第一 DMA 描述符数据和第二 DMA 描述符数据, 所述第一 DMA 描述符命令包括用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址, 所述第一 DMA 描述符数据包括第一用于信息处理设备的地址以及第二用于缓冲存储器的地址, 所述第二 DMA 描述符数据包括第二用于信息处理设备的地址以及第三用于缓冲存储器的地址; 其中, 所述第一用于缓冲存储器的地址同所述第一存储单元相对应, 所述第二用于缓冲存储器的地址同所述第二存储单元相对应, 所述第三用于缓冲存储器的地址同所述第三存储单元相对应;

在所述存储设备和所述信息处理设备之间依据所述第一用于信息处理设备的地址和所述第二用于缓冲存储器的地址进行 DMA 传输, 以及依据所述第二用于信息处理设备的地址和所述第三用于缓冲存储器的地址进行 DMA 传输;

接收来自所述存储设备的消息, 所述消息指示所述存储设备对第一 DMA 描述符已执行完成;

释放所述第一存储单元、所述第二存储单元与所述第三存储单元。
8. 根据权利要求 7 所述的方法, 所述第一 IO 请求指示将所述第一用于信息处理设备的地址处的数据以及所述第二用于信息处理设备的地址处的数据写入到用于所述存储设备的闪存芯片的地址处。
9. 根据权利要求 7、8 中的方法, 其中分配第一存储单元、第二存储单元与第三存储

单元后，将所述第一存储单元、所述第二存储单元和所述第三存储单元链接为循环链表。

10. 根据权利要求 9 之一的方法，所述来自所述存储设备的消息中包括指示所述第一存储单元、所述第二存储单元和所述第三存储单元之一的内容。

11. 根据权利要求 1、3、7 之一的方法，其中分配所述第一存储单元后，在所述第一存储单元中存储用于标识所述第一 IO 请求的指针。

12. 一种在信息处理设备和存储设备之间进行数据传输的方法，所述存储设备包括缓冲存储器和闪存芯片，所述方法包括：

接收第一写请求，所述第一写请求包括要写入的数据和用于所述存储设备的闪存芯片的地址；

为所述第一写请求在所述信息处理设备中分配第一存储单元；

向所述存储设备发送第一写命令，所述第一写命令中包括数据、用于所述存储设备的闪存芯片的地址以及用于所述缓冲存储器的地址；其中，所述用于缓冲存储器的地址同所述第一存储单元相对应；

接收来自所述存储设备的消息，所述消息指示所述存储设备对所述写命令已执行完成；

释放所述第一存储单元。

13. 根据权利要求 12 所述的方法，其中还为所述第一写请求分配第二存储单元，并在所述第二存储单元中存储指向所述第一存储单元的指针；

基于所述消息，释放所述第一存储单元与所述第二存储单元。

14. 一种在信息处理设备和存储设备之间进行 DMA 传输的装置，所述存储设备包括缓冲存储器和闪存芯片，所述装置包括：

用于接收第一 IO 请求的模块；

用于为所述第一 IO 请求分配第一存储单元与第二存储单元的模块；

用于向所述存储设备发送所述第一 DMA 描述符的模块，所述第一 DMA 描述符包括用于信息处理设备的地址、用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址与第二用于缓冲存储器的地址，其中，所述第一用于缓冲存储器的地址同所述第一存储单元相对应，所述第二用于缓冲存储器的地址同所述第二存储单元相对应；

用于在所述存储设备和所述信息处理设备之间依据所述第一 DMA 描述符进行 DMA 传输的模块；

用于接收来自所述存储设备的消息的模块，所述消息指示所述存储设备对第一 DMA 描述符已执行完成；

用于释放所述第一存储单元与所述第二存储单元的模块。

15. 一种在信息处理设备和存储设备之间进行 DMA 传输的装置，所述存储设备包括缓冲存储器和闪存芯片，所述装置包括：

用于接收第一 IO 请求的模块；

用于为所述第一 IO 请求分配第一存储单元与第二存储单元的模块；

用于向所述存储设备发送所述第一 DMA 描述符的模块，所述第一 DMA 描述符包括第一 DMA 描述符命令和第一 DMA 描述符数据，所述第一 DMA 描述符命令包括用于所述存储设备的闪存芯片的地址以及第一用于缓冲存储器的地址，所述第一 DMA 描述符数据包括用于信息处理设备的地址以及第二用于缓冲存储器的地址；其中，所述第一用于缓冲存储器的地址同所述第一存储单元相对应，所述第二用于缓冲存储器的地址同所述第二存储单元相对应；

用于在所述存储设备和所述信息处理设备之间依据所述用于信息处理设备的地址和所述第二用于缓冲存储器的地址进行 DMA 传输的模块；

用于接收来自所述存储设备的消息的模块，所述消息指示所述存储设备对第一 DMA 描述符已执行完成；

用于释放所述第一存储单元与所述第二存储单元的模块。

16. 一种在信息处理设备和存储设备之间进行数据传输的装置，所述存储设备包括缓冲存储器和闪存芯片，所述装置包括：用于接收第一写请求的装置，所述第一写请求包括要写入的数据和用于所述存储设备的闪存芯片的地址；用于为所述第一写请求在所述信息处理设备中分配第一存储单元的装置；用于向所述存储设备发送第一写命令的装置，所述第一写命令中包括数据、用于所述存储设备的闪存芯片的地址以及用于所述缓冲存储器的地址；其中，所述用于缓冲存储器的地址同所述第一存储单元相对应；用于接收来自所述存储设备的消息的装置，所述消息指示所述存储设备对所述写命令已执行完成；用于释放所述第一存储单元的装置。

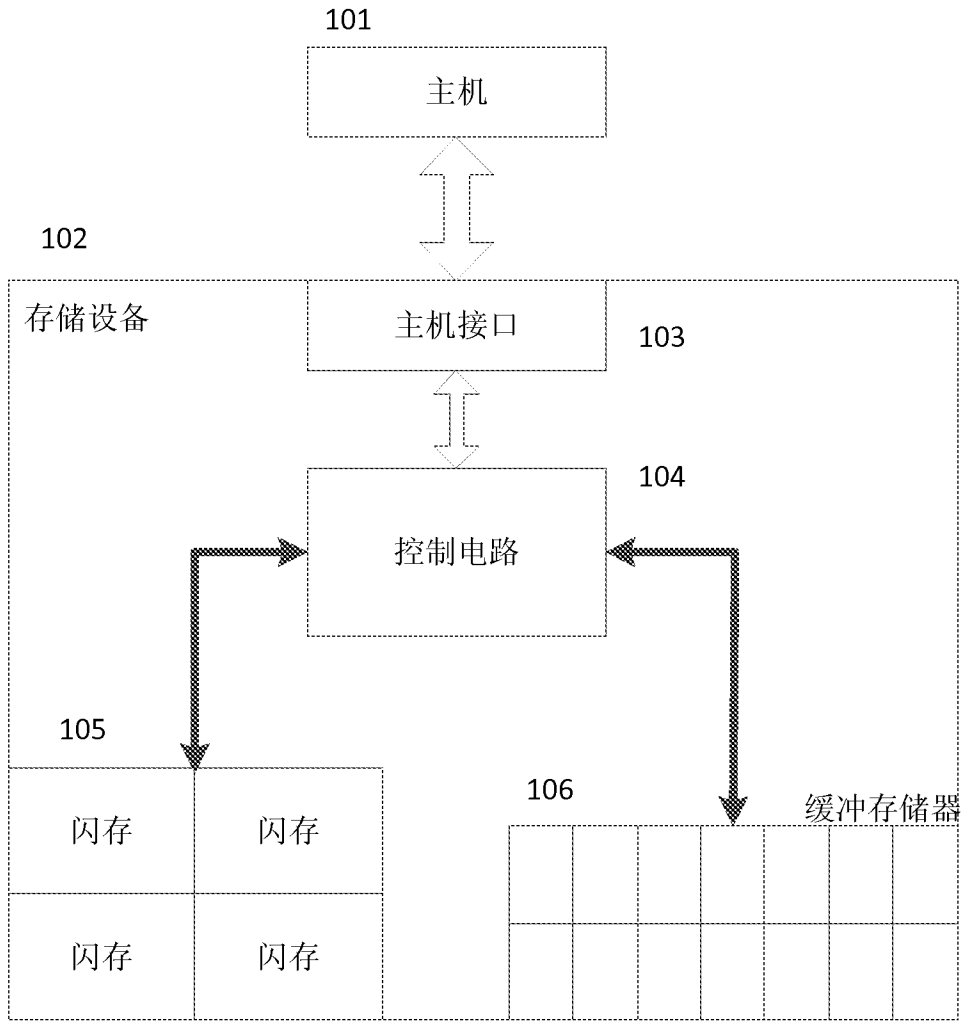


图 1

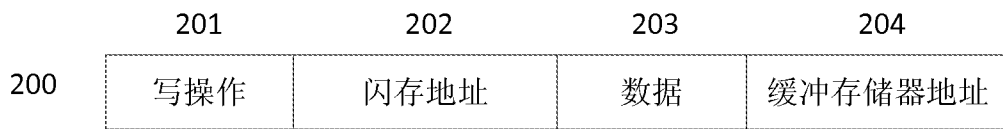


图2A

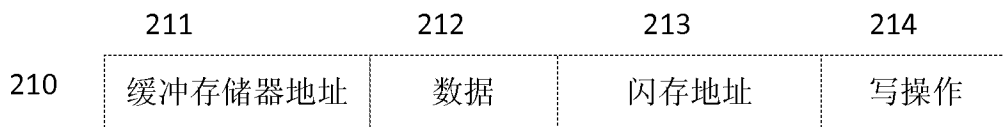


图2B

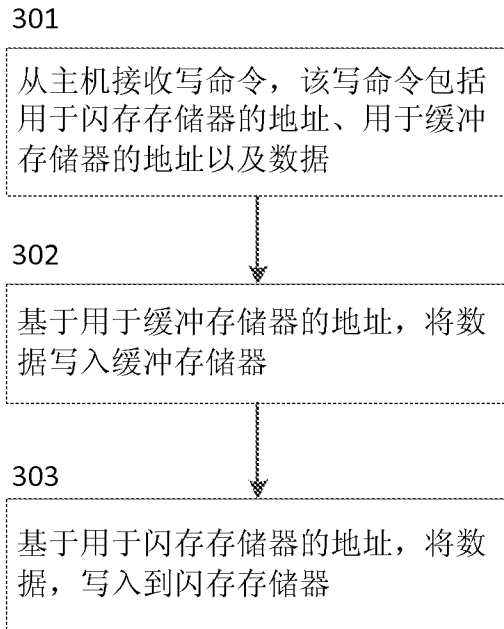


图 3

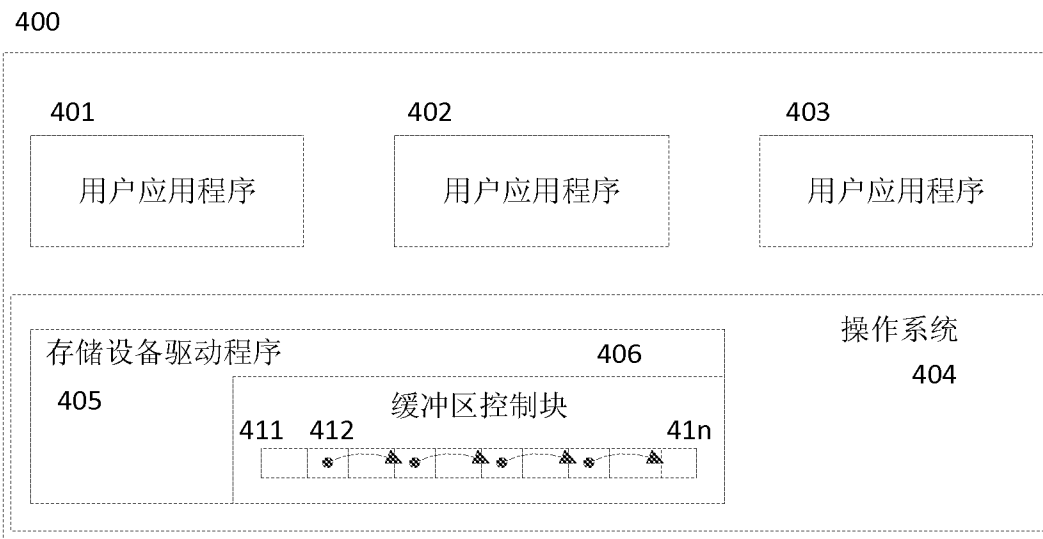


图 4

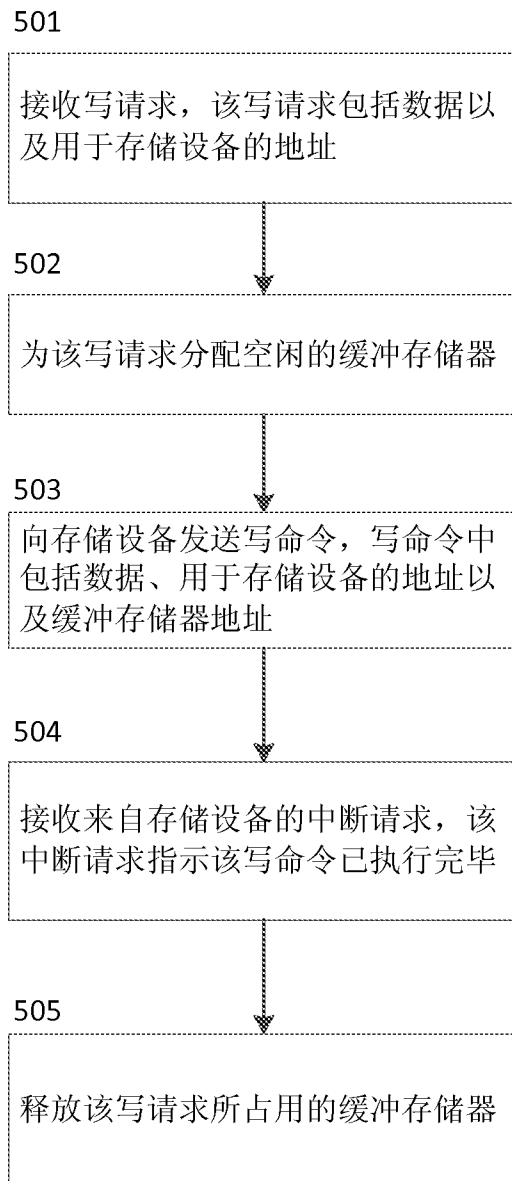


图 5

600

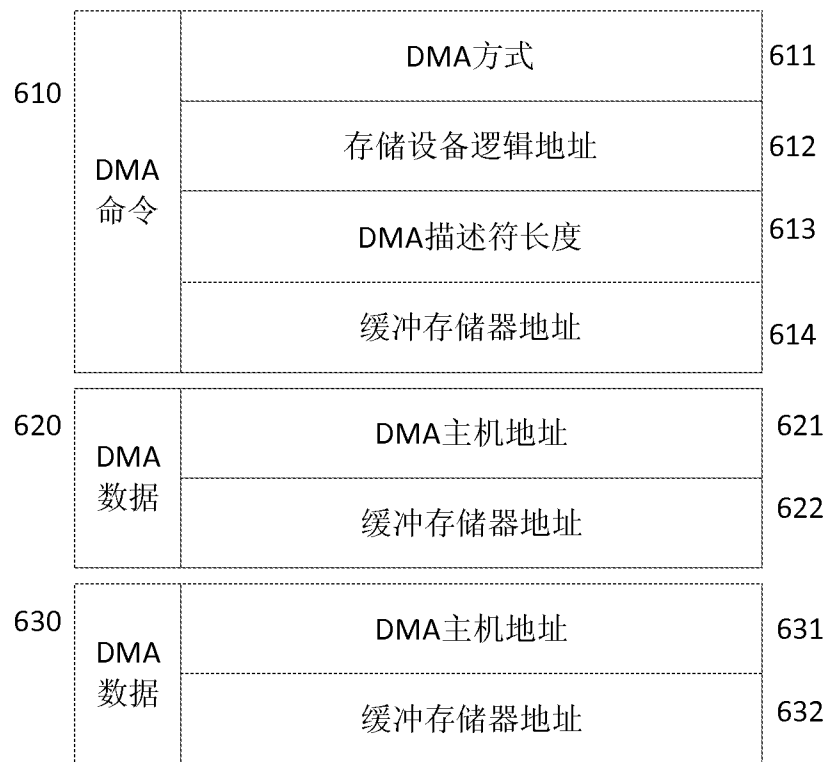


图 6

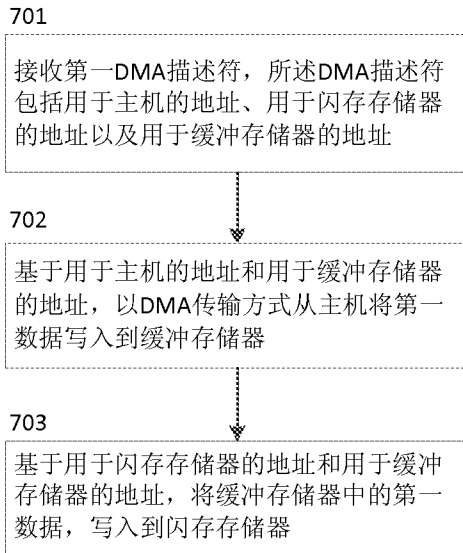


图7A

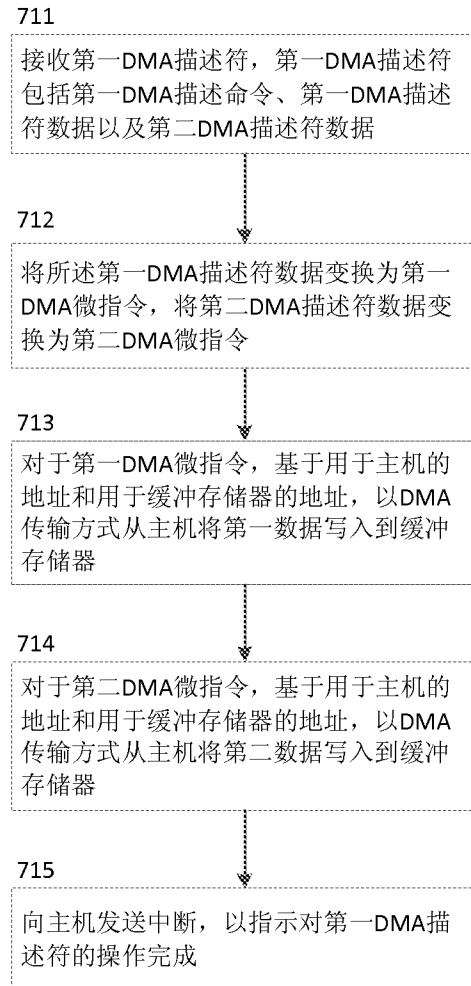


图7B

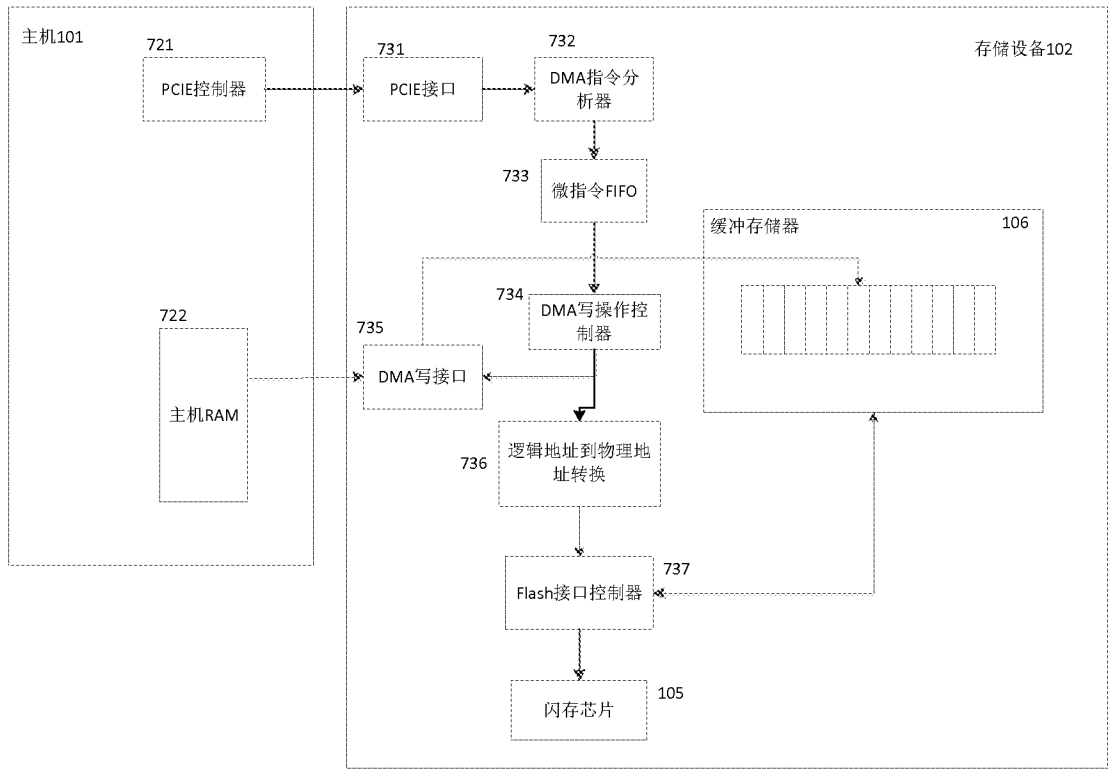


图 7C

800

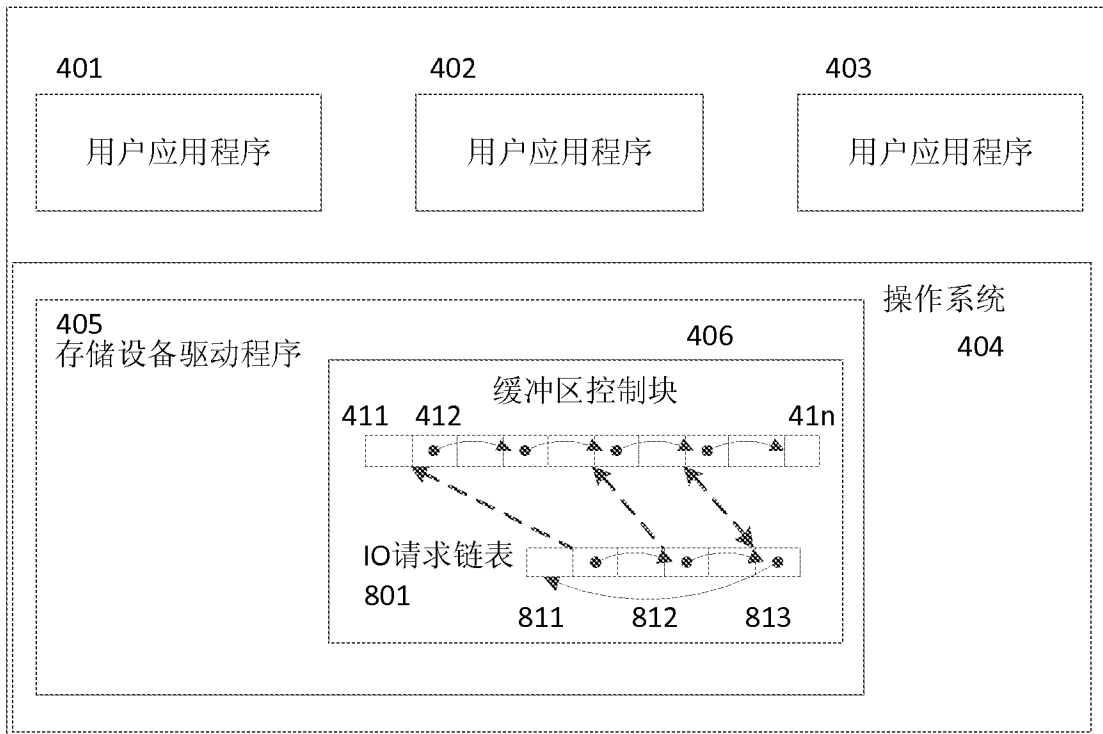


图 8

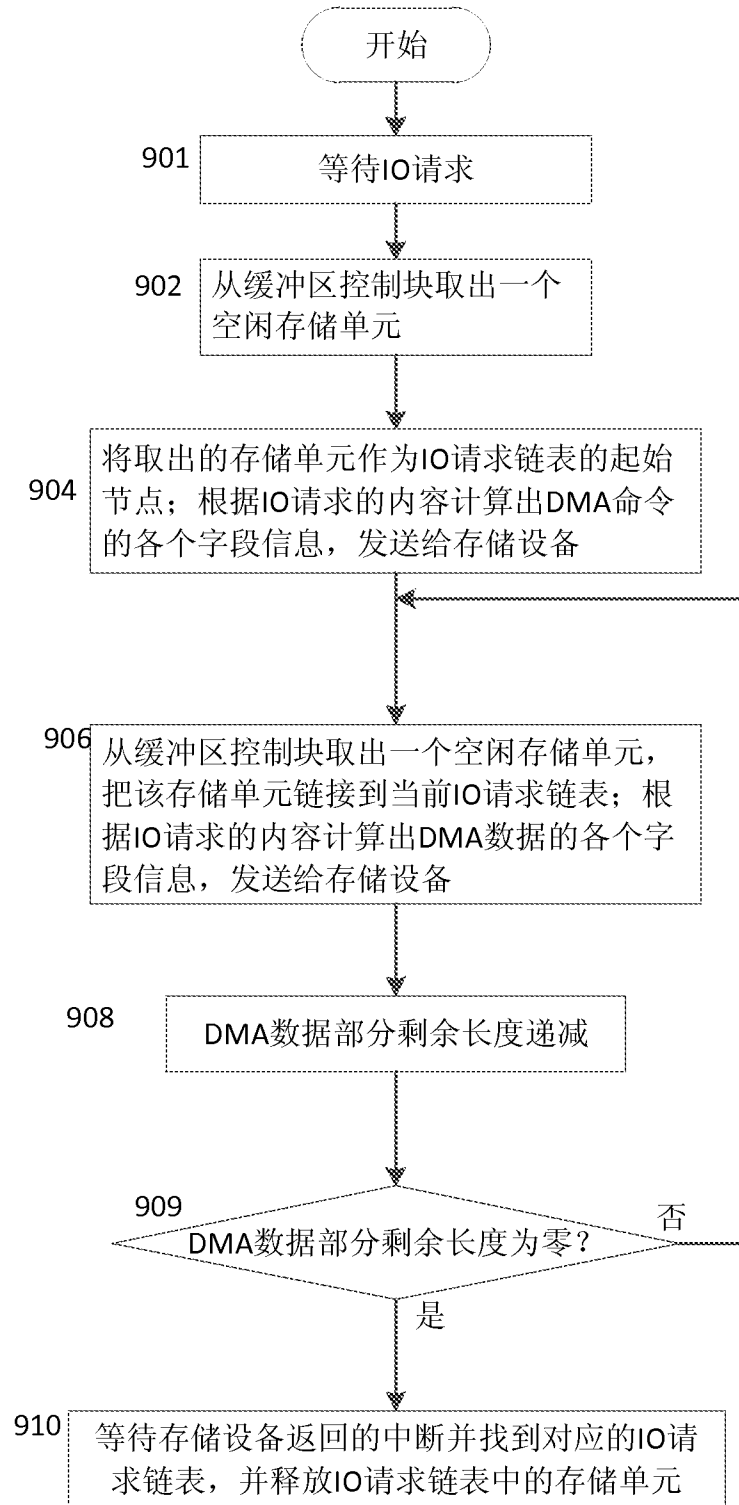


图 9A

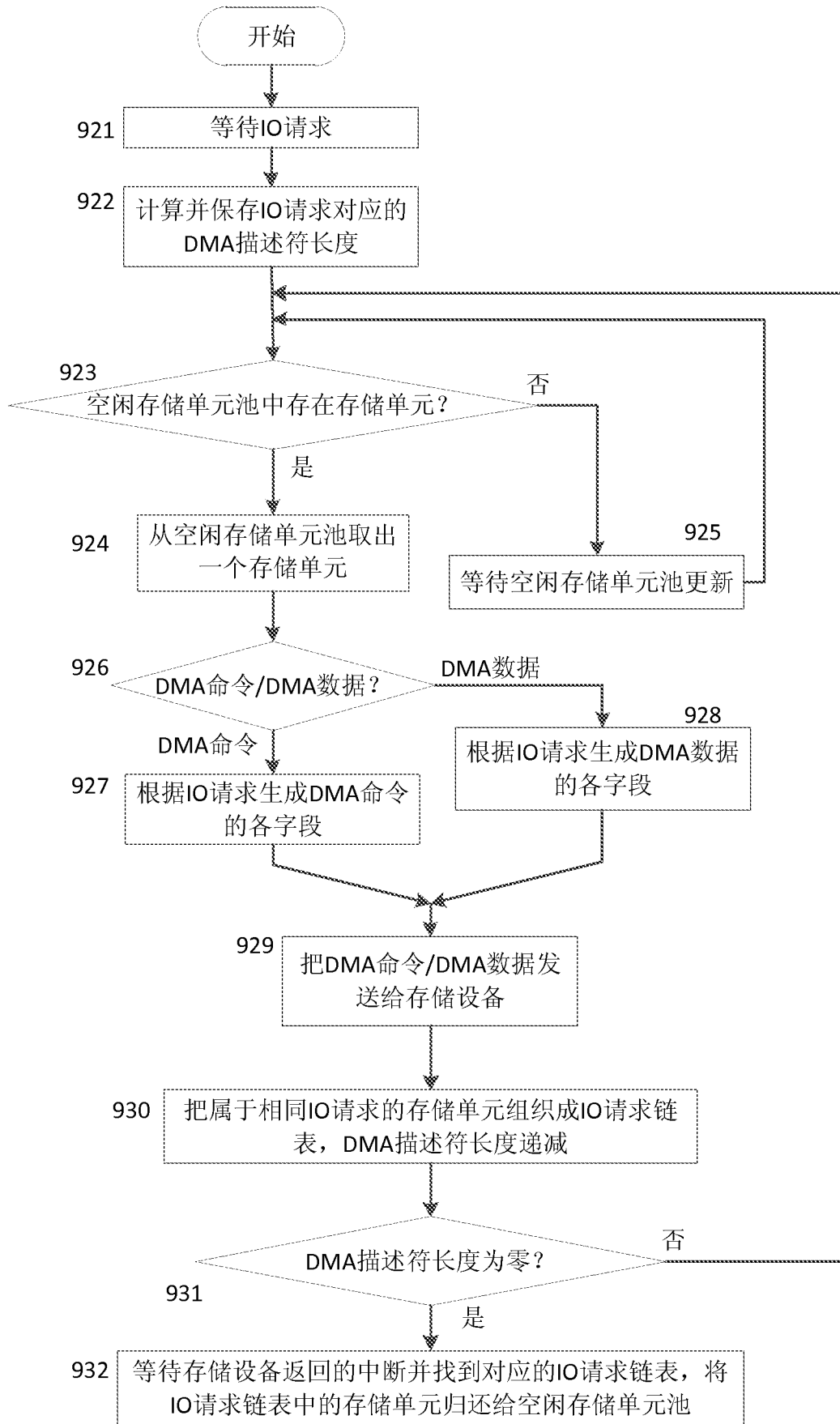


图 9B

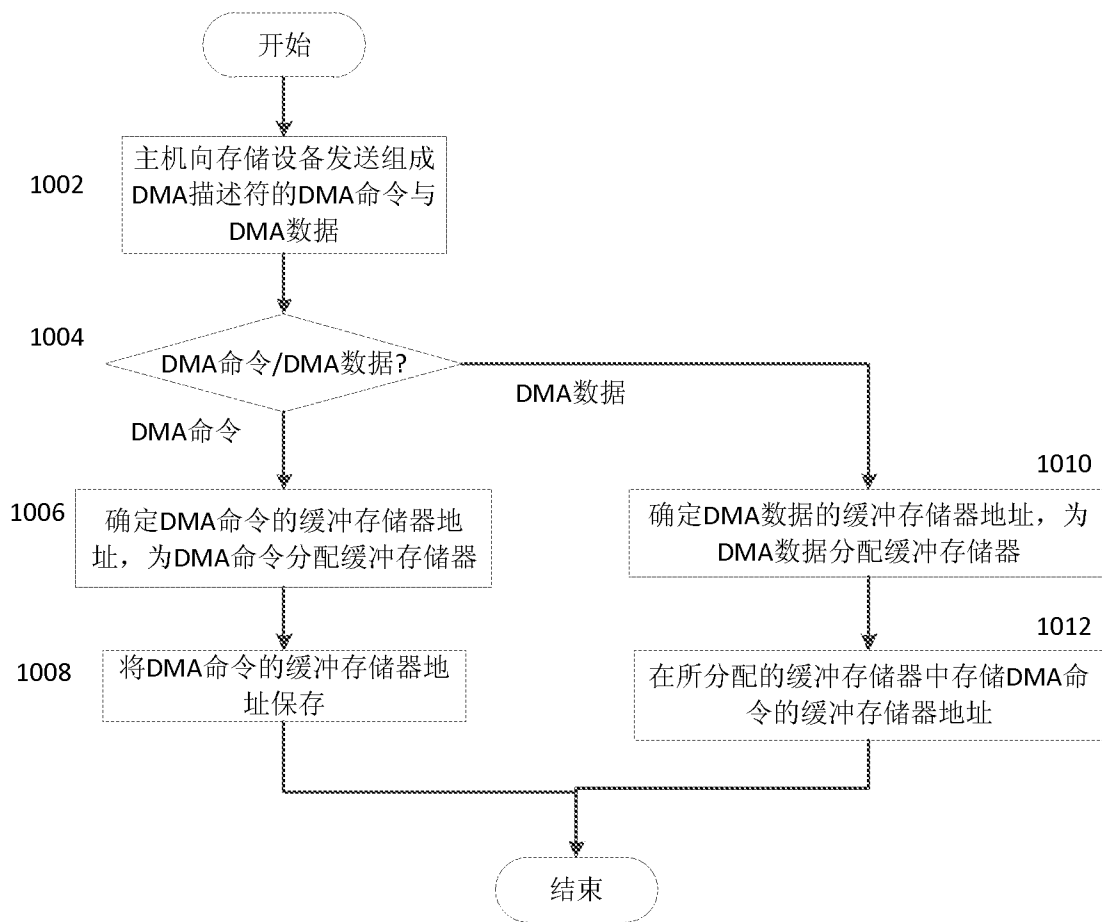


图10A

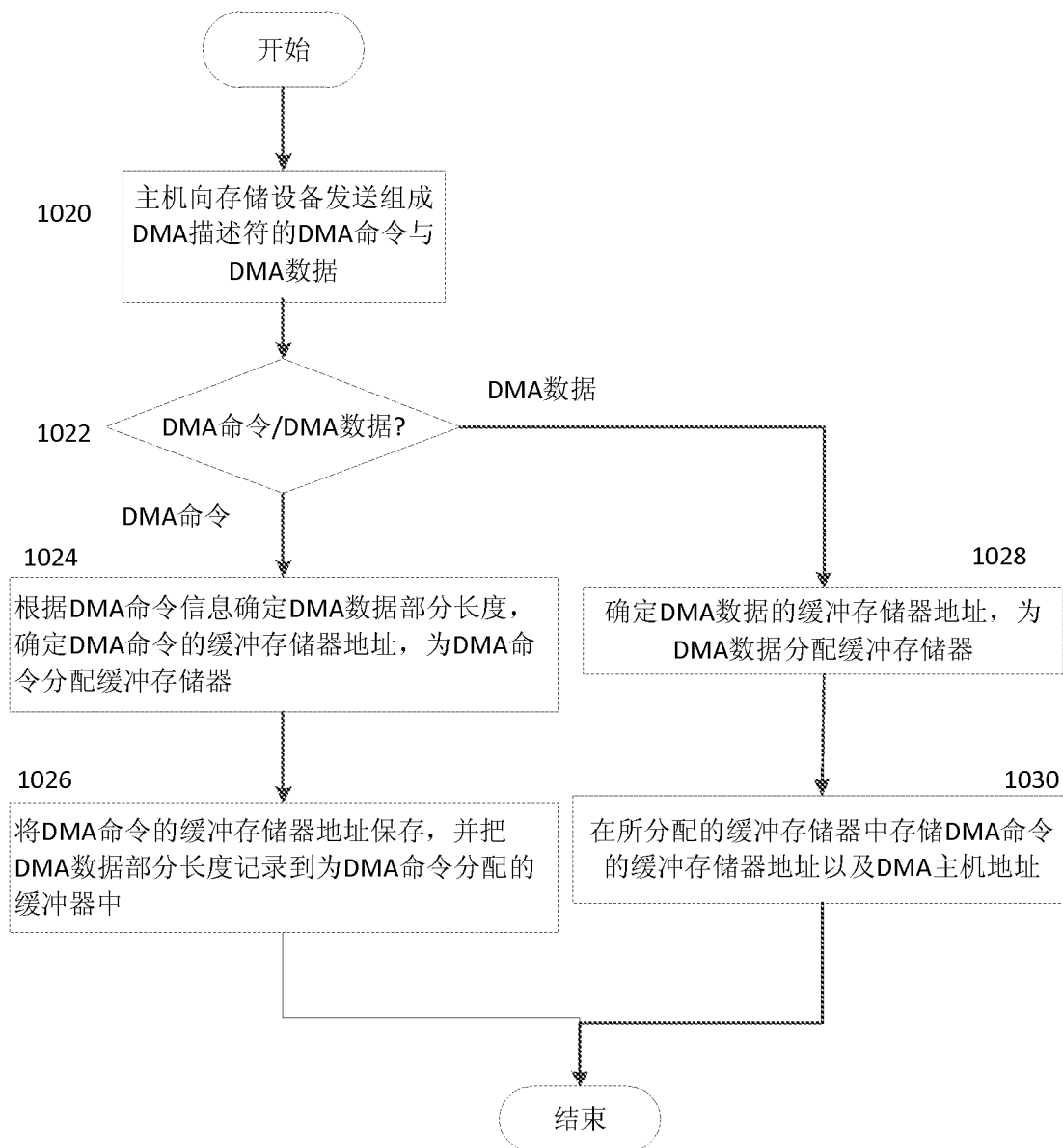


图 10B

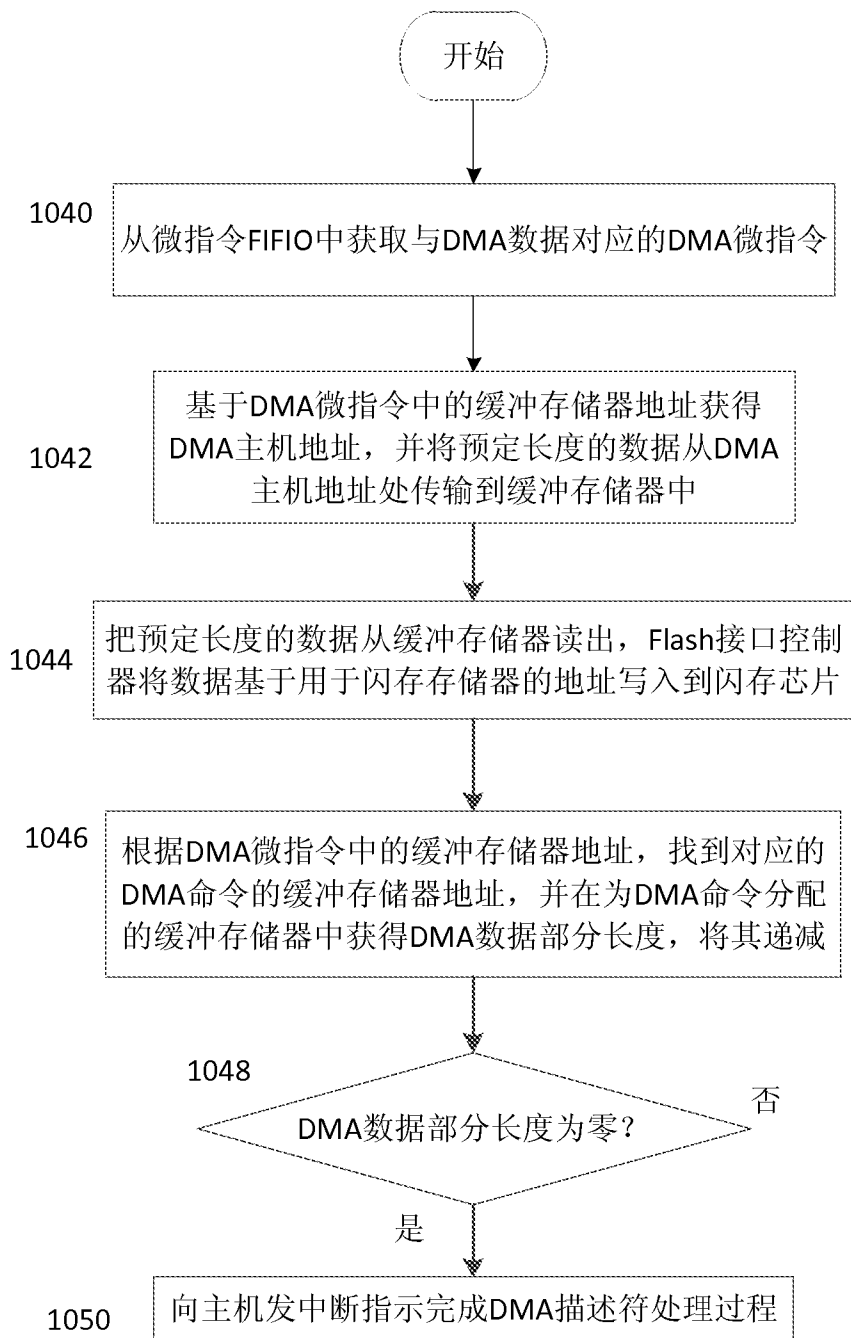


图10C

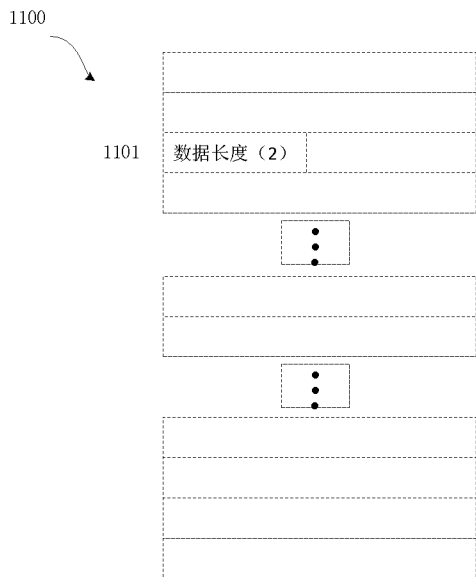


图11A

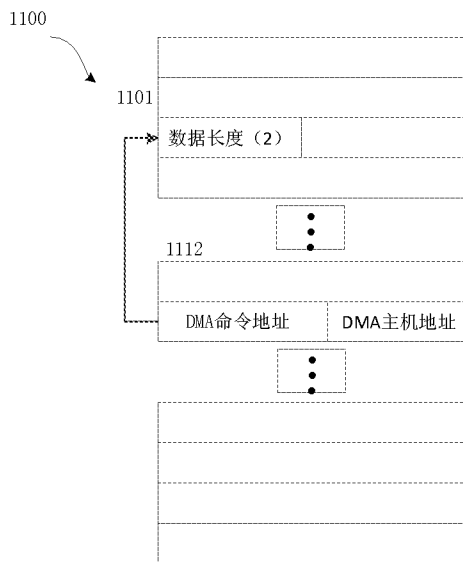


图11B

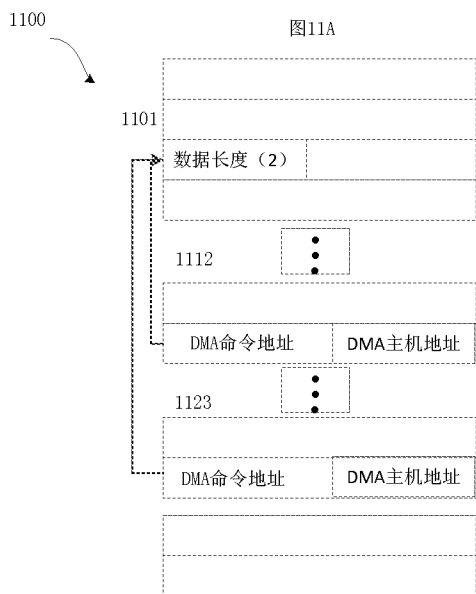


图11C

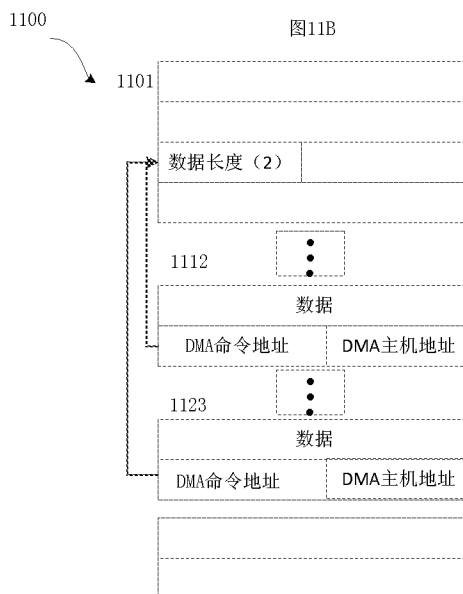


图11D

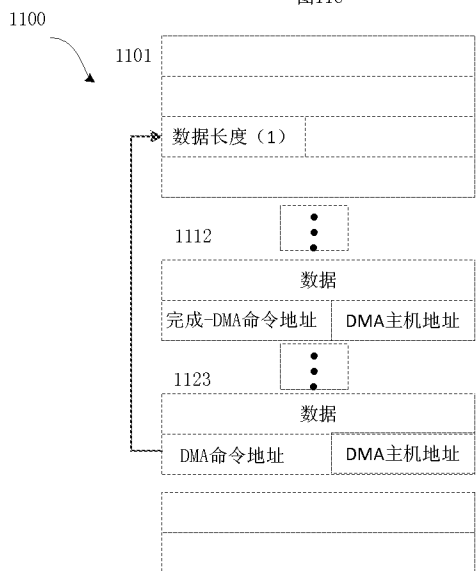


图11E

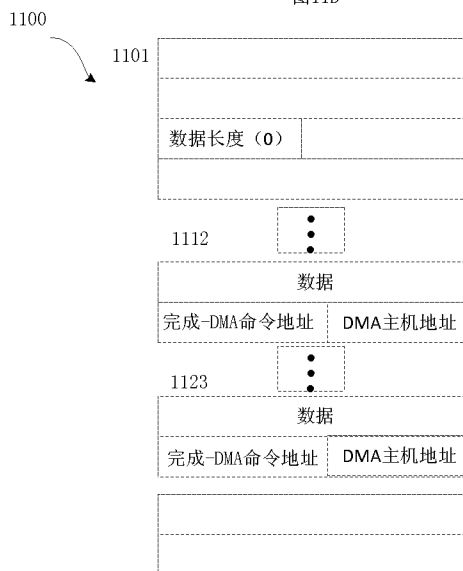


图11F

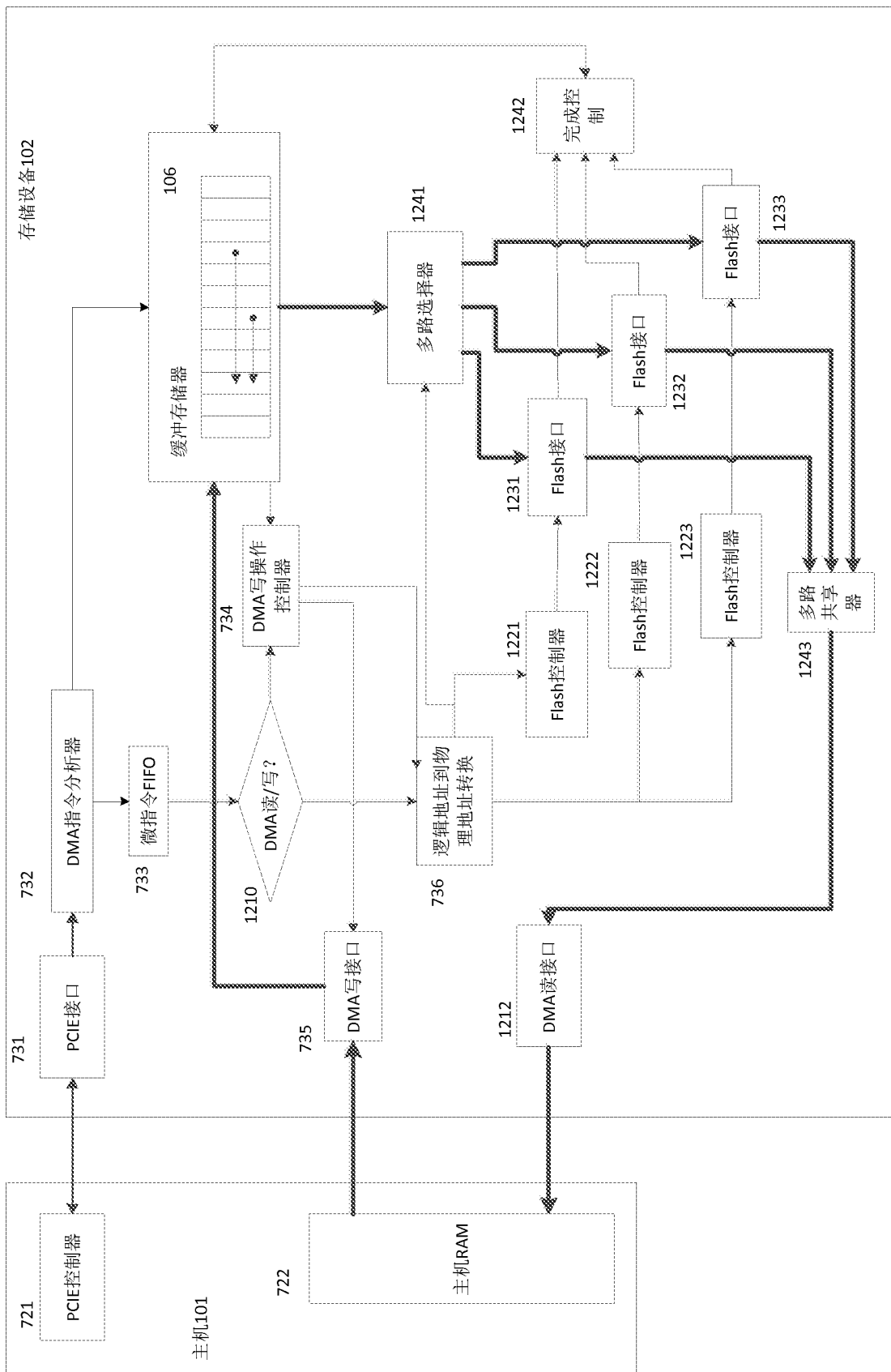


图 12

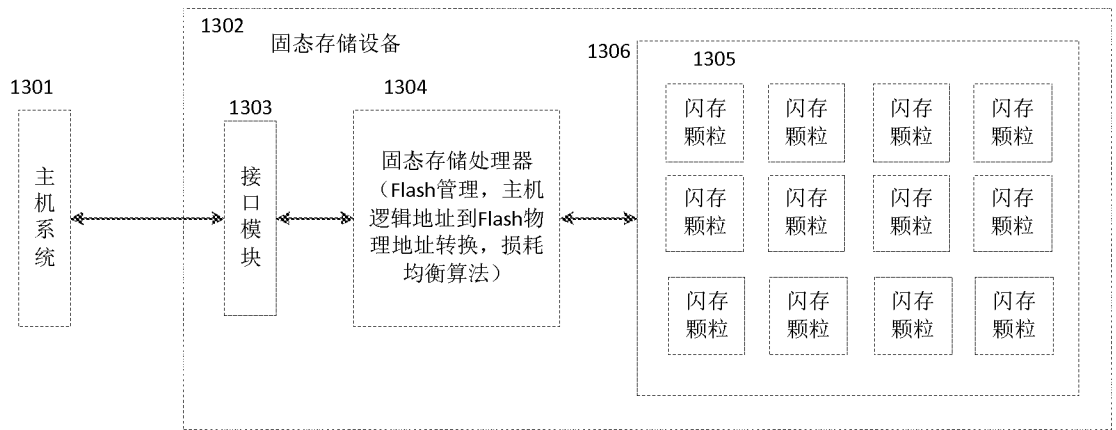


图 13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/075523

A. CLASSIFICATION OF SUBJECT MATTER

G06F 13/28 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNTXT, WPI, EPODOC: DMA, flash, memory, SSD, buffer, address, instruction, request, order, injunction, writ+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X	CN 102693198 A (BEIJING YIHENG CHUANGYUAN TECHNOLOGY CO) 26 September 2012 (26.09.2012) see claims 1 to 16	1-16
Y	CN 1527210 A (SANDISK IL LTD) 08 September 2004 (08.09.2004) see description, page 12, line 18 to page 13, line 6, and figure 9	12-13, 16
Y	CN 102150131 A (MICROSOFT CORP.) 10 August 2011 (10.08.2011) see claims 1 to 8	12-13, 16
A	CN 101354906 A (HUNAN YUANKE INNOVATION TECHNOLOGY CO LTD) 28 January 2009 (28.01.2009) see the whole document	1-16

Further documents are listed in the continuation of Box C. See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search
07 August 2013 (07.08.2013)

Date of mailing of the international search report
22 August 2013 (22.08.2013)

Name and mailing address of the ISA
State Intellectual Property Office of the P. R. China
No. 6, Xitucheng Road, Jimenqiao
Haidian District, Beijing 100088, China
Facsimile No. (86-10) 62019451

Authorized officer
FENG, Huiping
Telephone No. (86-10) 62411838

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/075523

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102693198 A	26.09.2012	None	
CN 1527210 A	08.09.2004	WO 0060476 A1	12.10.2000
		AU 3756400 A	23.10.2000
		US 6148354 A	14.11.2000
		BR 0006063 A	20.03.2001
		EP 1092193 A1	18.04.2001
		CN 1304509 A	18.07.2001
		KR 20010071332 A	28.07.2001
		JP 2002541554 A	03.12.2002
		AU 766478 B	16.10.2003
		TW 550454 A	01.09.2003
		KR 20030084947 A	01.11.2003
		IL 139662 A	28.03.2004
		AU 2003268851 A1	22.01.2004
		EP 1092193 B1	11.05.2005
		DE 60020046 E	16.06.2005
		EP 1548604 A2	29.06.2005
		ES 2241593 T3	01.11.2005
		AU 2003268851 B2	24.11.2005
		JP 2006031733 A	02.02.2006
		SG 117466 A1	29.12.2005
		DE 60020046 T2	26.01.2006
		INDEL 200401900 A	28.07.2006
		KR 100505972 B	04.08.2005
		AU 2006200756 A1	16.03.2006
		CN 1264100 C	12.07.2006
		EP 1746513 A2	24.01.2007
		IL 158578 A	10.12.2006
		DE 20023887 U1	22.03.2007
		SG 131813 A1	28.05.2007
		CN 1937073 A	28.03.2007
		JP 2007200351 A	09.08.2007
		KR 20070015480 A	02.02.2007
		CN 100385426 C	30.04.2008
		AU 2006200756 B2	03.04.2008

(NEXTTOSEETHEEXTRASH)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/075523

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
		AU 2008202866 A1	24.07.2008
		EP 1746513 A3	26.11.2008
		CN 101345077 A	14.01.2009
		INDEL 200401899 A	26.12.2008
		KR 20080098450 A	07.11.2008
		JP 4261069 B2	30.04.2009
		EP 1092193 B2	16.09.2009
		KR 100914427 B1	28.08.2009
		EP 2120435 A2	18.11.2009
		EP 1548604 B1	30.12.2009
		DE 60043623 E	11.02.2010
		KR 100922766 B1	21.10.2009
		EP 2163991 A2	17.03.2010
		EP 2120435 A3	05.05.2010
		EP 2163991 A3	05.05.2010
		EP 1746513 B1	05.05.2010
		DE 60044381 E	17.06.2010
		SG 163430 A1	30.08.2010
		AU 2008202866 A1	24.07.2008
		AU 2010257369 A1	20.01.2011
		AU 2008202866 B2	03.03.2011
		JP 2011054187 A	17.03.2011
		USRE 42397 E	24.05.2011
		IN 195845 B	26.05.2006
		IN 251251 B	09.03.2012
		JP 5044254 B2	10.10.2012
		AU 2010257369 B2	20.09.2012
		AU 2012216828 A1	04.10.2012
		SG 186496 A1	30.01.2013
		EP 2120435 B1	24.04.2013
		EP 2163991 B1	24.04.2013
CN 102150131 A	10.08.2011	WO 2010030715 A2	18.03.2010
		US 2010070701 A1	18.03.2010
		US 2010070747 A1	18.03.2010
		WO 2010030715 A3	06.05.2010

(Next to see the extra sheet II)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/075523

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
		TW 201011538 A	16.03.2010
		EP 2329360 A2	08.06.2011
		US 8032707 B2	04.10.2011
		INDELNP 201101942 E	16.12.2011
		US 2011314202 A1	22.12.2011
		JP 2012503232 A	02.02.2012
		US 8135914 B2	13.03.2012
		US 2012173824 A1	05.07.2012
CN 101354906 A	28.01.2009	CN 101354906 B	15.09.2010

国际检索报告

国际申请号
PCT/CN2013/075523

A. 主题的分类		
G06F 13/28 (2006.01) i		
按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类		
B. 检索领域		
检索的最低限度文献(标明分类系统和分类号)		
IPC: G06F		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))		
CNABS, CNTXT, WPI, EPODOC: DMA、直接访问存储器、闪存、缓冲存储器、缓存、地址、指令、命令、请求、写、编程、flash、memory、SSD、buffer、address、instruction、request、order、injunction、writ+		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
P, X	CN102693198A (北京亿恒创源科技有限公司) 26.9 月 2012 (26.09.2012) 权利要求 1-16	1-16
Y	CN1527210A (M—系统快闪盘开拓者公司) 08.9 月 2004 (08.09.2004) 参见说明书第 12 页第 18 行-第 13 页第 6 行、图 9	12-13, 16
Y	CN102150131A (微软公司) 10.8 月 2011 (10.08.2011) 参见权利要求 1-8	12-13, 16
A	CN101354906A (湖南源科创新科技股份有限公司) 28.1 月 2009 (28.01.2009) 参见全文	1-16
<input type="checkbox"/> 其余文件在 C 栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期 07.8 月 2013(07.08.2013)		国际检索报告邮寄日期 22.8 月 2013 (22.08.2013)
ISA/CN 的名称和邮寄地址: 中华人民共和国国家知识产权局 中国北京市海淀区蓟门桥西土城路 6 号 100088 传真号: (86-10)62019451		受权官员 冯慧萍 电话号码: (86-10) 62411838

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2013/075523

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
CN102693198A	26.09.2012	无	
CN1527210A	08.09.2004	WO0060476A1	12.10.2000
		AU3756400A	23.10.2000
		US6148354A	14.11.2000
		BR0006063A	20.03.2001
		EP1092193A1	18.04.2001
		CN1304509A	18.07.2001
		KR20010071332A	28.07.2001
		JP2002541554A	03.12.2002
		AU766478B	16.10.2003
		TW550454A	01.09.2003
		KR20030084947A	01.11.2003
		IL139662A	28.03.2004
		AU2003268851A1	22.01.2004
		EP1092193B1	11.05.2005
		DE60020046E	16.06.2005
		EP1548604A2	29.06.2005
		ES2241593T3	01.11.2005
		AU2003268851B2	24.11.2005
		JP2006031733A	02.02.2006
		SG117466A1	29.12.2005
		DE60020046T2	26.01.2006
		INDEL200401900A	28.07.2006
		KR100505972B	04.08.2005
		AU2006200756A1	16.03.2006
		CN1264100C	12.07.2006
		EP1746513A2	24.01.2007
		IL158578A	10.12.2006
		DE20023887U1	22.03.2007
		SG131813A1	28.05.2007
		CN1937073A	28.03.2007
		JP2007200351A	09.08.2007
		KR20070015480A	02.02.2007
		CN100385426C	30.04.2008
(下接附加页 1)		AU2006200756B2	03.04.2008

国际检索报告

国际申请号
PCT/CN2013/075523

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
		AU2008202866A1	24.07.2008
		EP1746513A3	26.11.2008
		CN101345077A	14.01.2009
		INDEL200401899A	26.12.2008
		KR20080098450A	07.11.2008
		JP4261069B2	30.04.2009
		EP1092193B2	16.09.2009
		KR100914427B1	28.08.2009
		EP2120435A2	18.11.2009
		EP1548604B1	30.12.2009
		DE60043623E	11.02.2010
		KR100922766B1	21.10.2009
		EP2163991A2	17.03.2010
		EP2120435A3	05.05.2010
		EP2163991A3	05.05.2010
		EP1746513B1	05.05.2010
		DE60044381E	17.06.2010
		SG163430A1	30.08.2010
		AU2008202866A1	24.07.2008
		AU2010257369A1	20.01.2011
		AU2008202866B2	03.03.2011
		JP2011054187A	17.03.2011
		USRE42397E	24.05.2011
		IN195845B	26.05.2006
		IN251251B	09.03.2012
		JP5044254B2	10.10.2012
		AU2010257369B2	20.09.2012
		AU2012216828A1	04.10.2012
		SG186496A1	30.01.2013
		EP2120435B1	24.04.2013
		EP2163991B1	24.04.2013
CN102150131A	10.08.2011	WO2010030715A2	18.03.2010
		US2010070701A1	18.03.2010
		US2010070747A1	18.03.2010
		WO2010030715A3	06.05.2010

(下接附加页 2)

国际检索报告

国际申请号
PCT/CN2013/075523

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
		TW201011538A	16.03.2010
		EP2329360A2	08.06.2011
		US8032707B2	04.10.2011
		INDELNP201101942E	16.12.2011
		US2011314202A1	22.12.2011
		JP2012503232A	02.02.2012
		US8135914B2	13.03.2012
		US2012173824A1	05.07.2012
CN101354906A	28.01.2009	CN101354906B	15.09.2010