



## (12) 发明专利申请

(10) 申请公布号 CN 102622316 A

(43) 申请公布日 2012. 08. 01

(21) 申请号 201110448623. 1

(22) 申请日 2011. 12. 22

(30) 优先权数据

12/976, 608 2010. 12. 22 US

(71) 申请人 西部数据技术公司

地址 美国加利福尼亚州

(72) 发明人 J·布思 R·A·威尔森

(74) 专利代理机构 北京纪凯知识产权代理有限公司  
11245

代理人 赵蓉民

(51) Int. Cl.

G06F 13/16 (2006. 01)

权利要求书 2 页 说明书 5 页 附图 4 页

(54) 发明名称

选择性地使能主机传递中断

(57) 摘要

本发明涉及选择性地使能主机传递中断。本发明的实施例涉及减少非易失性存储装置的控制器上的中断数量以提高存储系统的数据传递性能的系统和方法。此处描述的实施例选择性地使能主机传递硬件产生的针对主机命令的中断。通过考虑命令类型、接口资源对接受额外主机传递的可用性和命令大小,所述中断可被使能或者禁止。此处描述的实施例对于用有限范围的标识标签实现用于主机传递的加标签方案的主机接口有用。

1. 一种能够减少中开销的非易失性存储系统,所述非易失性存储系统包括:

接口,所述接口被配置以从主机装置接收命令,所述命令包括标识标签,所述标识标签从标识标签的有限集中选择,每个标识标签对于未完成的命令是唯一的,所述接口附加地被配置以使用主机通信线程调度主机数据传递,所述主机通信线程能够在主机传递完成时产生中断;以及

处理器,所述处理器被配置以管理通过所述接口与所述主机装置的通信,以及在主机传递完成时选择性地禁止所述接口产生中断,所述选择性的禁止至少部分基于命令类型、接口资源接受额外主机传递的可用性和命令大小中之一,由此所述选择性的禁止减少与执行所述命令相关联的开销。

2. 根据权利要求 1 所述的非易失性存储系统,其中所述处理器选择性地禁止为读命令产生中断。

3. 根据权利要求 1 所述的非易失性存储系统,其中所述非易失性存储系统包括固态非易失性存储器阵列。

4. 根据权利要求 1 所述的非易失性存储系统,其中所述标识标签包括加标签的命令排队标签和原生命令排队标签之一。

5. 根据权利要求 1 所述的非易失性存储系统,其中所述接口使用串行 ATA 接口与所述主机装置通信。

6. 根据权利要求 1 所述的非易失性存储系统,其中所述接口使用串行附接 SCSI 接口与所述主机装置通信。

7. 根据权利要求 1 所述的非易失性存储系统,其中所述标识标签由特定协议分配。

8. 根据权利要求 1 所述的非易失性存储系统,其中所述处理器还被配置以通过释放分配给主机数据传递完成的命令的资源对中断做出响应。

9. 根据权利要求 8 所述的非易失性存储系统,其中所述处理器还被配置以当所述中断被选择性地禁止时重用分配给所述命令的资源。

10. 根据权利要求 9 所述的非易失性存储系统,其中所述资源被另一个命令重用,该另一个命令包括与所述命令包括的标识标签相同的标识标签。

11. 根据权利要求 1 所述的非易失性存储系统,其中根据每个主机传递的完成,所述接口产生的中断被可配置地产生。

12. 根据权利要求 1 所述的非易失性存储系统,其中根据一批主机传递的完成,所述接口产生的中断被可配置地产生。

13. 一种主机传递请求处理的方法,所述方法包括:

通过与主机系统通信的非易失性存储装置:

从所述主机系统接收主机传递请求,所述主机传递请求与从标识标签的有限集中选择的标识标签相关联,每个标识标签对于未完成的命令是唯一的;

使用主机接口执行所述主机传递请求;以及

当主机传递请求完成时确定是否中断非易失性存储装置的处理器,所述确定是否中断处理器是至少部分基于至少以下一个:所分配的所述与主机传递请求相关联的资源和接口资源接受额外主机传递请求的可用性。

14. 根据权利要求 13 所述的主机传递请求处理的方法,还包括针对主机传递请求的使

用分配资源,所分配的资源与所述标识标签的有限集中的每个标识标签相关联,其中分配步骤在接收步骤之前进行。

15. 一种由实现原生命令排队的存储子系统执行的方法,所述方法包括:

从主机系统接收加标签的读命令,以及在所述存储子系统的队列中存储所述加标签的读命令;

执行来自所述队列的所述加标签的读命令,其中执行所述加标签的读命令包括向所述主机系统传递由所述存储子系统存储的数据;以及

至少部分地基于是否任何其它命令存储在所述队列中确定是否执行与所述加标签的读命令相关联的中断处理。

16. 根据权利要求 15 所述的方法,其中所述方法包括作为对确定所述队列中没有存储其它命令的响应而禁止所述中断处理,由此避免与所述中断处理相关联的开销。

17. 根据权利要求 15 所述的方法,其中所述存储子系统根据 SATA 接口接收和执行所述加标签的读命令。

18. 根据权利要求 15 所述的方法,其中所述存储子系统根据 SAS 接口接收和执行所述加标签的读命令。

## 选择性地使能主机传递中断

### 技术领域

### 背景技术

[0001] 非易失性介质存储系统可从主机接收很多同时的命令。在特定的主机传递协议中,向介质存储系统发送具有唯一命令标识标签的命令。当存储系统从主机接收命令时,存储系统通常分配资源并且用信号通知主机以执行主机传递。在完成主机传递之后,存储系统的主机传递硬件启动对存储系统控制器的中断,其解分配或释放资源并且确定是否可开始新的主机传递。中断的处理可延迟存储系统控制器中的控制器执行其它处理。

### 发明内容

### 附图说明

[0002] 下面将参照附图描述实现本发明的各个特征的系统和方法,其中:

[0003] 图 1 示出本发明的实施例的系统概览。

[0004] 图 2 示出指示根据一个实施例与传递硬件通信的存储控制器的操作的流程图。

[0005] 图 3 示出指示根据一个实施例与具有减少的控制器中断的传递硬件通信的存储控制器的操作的流程图。

[0006] 图 4 示出描绘根据一个实施例的确定是否使能硬件传递中断的处理的流程图。

### 具体实施方式

[0007] 尽管描述了本发明的特定实施例,但是这些实施例仅作为示例呈现,而不旨在限制本发明的范围。实际上,此处描述的新颖方法和系统可以在各种其它形式中实现。此外,此处描述的方法和系统可以进行各种省略、替代和修改而不背离本发明的实质。所附的权利要求及其等同物旨在覆盖这些形式或者修改,因为其将落入本发明的范围和实质。

#### [0008] 系统概述

[0009] 本发明的实施例涉及减少非易失性存储系统上的控制器处理的中断以提高存储系统的数据传递性能的系统和方法。此处描述的实施例减少当往来主机的数据传递(“主机传递”)完成时候产生中断的情形。在一些实施例中,非易失性存储系统从主机接收具有标识标签的命令。在这些实施例中,可针对预限定范围的标识标签的主机传递分配资源,并且在特定情形下可安全地重用而没有中断,这减少每次主机传递对主机中断的需要。

[0010] 参照图 1,示出与主机装置 60 通信的非易失性存储装置 10 的实施例的系统图。在本实施例中,非易失性存储装置 10 包括非易失性内存存储器 20、装置控制器 30、命令队列 50、操作传递存储器 80 和硬件传递队列 90。主机装置 60 包括与命令队列 50 和硬件传递队列 90 通信的非易失性存储控制器 70。非易失性内存存储器 20 可包括固态存储器阵列、机械旋转磁盘阵列、包括固态和旋转磁盘阵列的混合体或者任何其它非易失性内存存储器类型。当主机装置向非易失性存储装置 10 发送命令时,该命令被接收在命令队列 50 中。在此

实施例中,主机装置 60 和非易失性存储装置 10 之间的接口发送包括标识标签的命令。在特定接口标准中,标识标签可用于允许多个未完成的请求在存储装置上挂起。例如,可针对原生命令排队配置 SATA 接口标准,其允许多达 32 个命令未完成并且使存储装置能够确定执行命令的顺序。其它标识标签标准包括加标签的命令排队。如图 1 所示,相同的标识标签与图 1 所示的各种队列内的存储器命令相关联,因而标识标签可用作公共基准。

[0011] 在一个实施例中,命令队列 50 从非易失性存储控制器 70 接收命令。命令可存储在命令队列 50 中直至被装置控制器 30 处理为止。当从命令队列 50 接收到命令时,装置控制器可以处理该命令接着准备主机传递。对于读命令,装置控制器 30 可以在向主机传递读数据之前存取非易失性内存存储器 20。对于写命令,装置控制器 30 可以请求主机传递以从主机获取要被写入内存存储器的数据。

[0012] 在一个实施例中,存储控制器可以从操作传递存储器 80 向主机传递请求分配缓冲器。主机传递请求可以然后被放置在主机传递队列 90 中,其可被硬件执行以完成主机队列请求从而向主机发送数据或者从主机接收数据。如果主机传递队列满,则在本实施例中装置控制器在分级队列 40 中存储多余的主机传递请求。

#### [0013] 主机传递完成中断

[0014] 图 2 是示出根据一个实施例的产生主机传递中断的处理的流程图。图 2 的标记为 100 的顶部指示可在存储装置控制器(诸如图 1 描绘的装置控制器 30)上执行的处理的部分。图 2 的标记为 110 的底部指示可在存储装置上的传递硬件上执行的处理的部分。

[0015] 在块 200,控制器接收准备用于主机传递的命令。在块 210,控制器检查传递队列是否满并且因此不能够接受额外的主机传递请求。如果主机传递队列满,则控制器可将命令放置在分级队列中,如块 220 指示的。命令将保留在分级队列中直至传递队列具有开口为止。否则,如果传递队列不满,则在块 230 命令可被放置在传递队列中。接着,传递硬件可在块 240 执行主机传递。在命令已经被放置在主机传递队列中之后,在主机传递挂起的同时控制器中的固件可继续完成其它任务,包括接收其它命令和将其它命令放置在传递队列中。

[0016] 在主机传递完成之后,在块 250 传递硬件产生中断以指示传递完成。当接收到中断时,在块 260 控制器 100 执行上下文切换以解决中断。在主机传递完成的同时控制器 100 可以执行其它处理,诸如存储器管理任务。这些处理被主机传递中断的完成中断。在中断之后,在块 270 控制器 100 释放向传递命令分配的资源(例如,操作传递存储器中的缓冲存储器的各部分)从而资源可被另一传递命令使用。在释放资源之后,在块 280 控制器 100 检查分级队列中是否存在命令。如果分级队列中有剩余命令,则通过返回到块 230 控制器可在传递队列中放置命令。如果分级队列中没有命令,则在块 290 处理结束并且在再次开始处理之前控制器等待新命令。

#### [0017] 减少主机传递中断

[0018] 如以上讨论的,在图 2 所示的实施例中,在每个传递完成时,传递硬件产生中断。这产生显著的装置控制器上的处理器被从其它管理任务中断的时间量。本发明的实施例提供一种替换方法,其降低使处理器中断的频率。通常,使处理器中断的需求是由对释放分配给刚完成的传递命令的资源的请求产生的。通过向系统提供操作传递存储器中的充分的缓冲器空间,存储装置可被配置为使得当完成针对控制器的每个主机传递时装置控制器不要

求中断,以管理分配用于处理主机传递的存储器资源。本发明的各种实施例利用以下事实,即对于主机和实现加标签的命令系统(诸如加标签的命令排队或者原生命令排队)的存储装置之间的接口,至存储装置的未完成命令最大可能数量被接口规范中标识标签的数量限制。两个这种接口是串行 ATA (SATA) 和串行附接 SCSI (SAS)。因此,在存储装置上提供充分的存储器资源以为全部标识标签同时提供主机传递资源可减少当中断时明确释放这些资源的需要。在一个实施例中,操作传递存储器的缓冲器中的特定存储器资源被分配到标识标签的每个受限范围,并且当具有相同标识标签的新命令被接收时资源被自动重用。尽管提供 SATA 和 SAS 作为示例,但是本发明的实施例不限于实现这些接口的存储装置,并且一般可应用于使用加标签的命令系统形式的任何实现方式。

[0019] 为了进一步说明各个实施例中使用的此配置,采用可被存储装置控制器处理的具有标识标签“5”的命令作为例子。在本例子中,存储装置已经为每个标识标签号码指定缓冲器地址,并且为标识标签“5”指定缓冲器地址“E”。在此设置中,当主机传递完成时装置控制器不需要被中断以明确释放缓冲器地址“E”。装置控制器理解,由于标识标签的受限范围,下一次具有相同标识标签“5”的命令被接收时,具有相同标识标签“5”的先前命令必须已经完成。因此,装置控制器针对新命令重用缓冲器地址“E”是安全的,而不依赖中断来通知它主机传递完成。

[0020] 参照图 3,提供示出根据一个实施例减少向控制器提供的中断数量的处理的流程图。和图 2 一样,图 3 的标记为 100 的顶部指示可在存储装置控制器(诸如图 1 描绘的装置控制器 30)上执行的处理的部分。图 3 的标记为 110 的底部指示可在存储装置上的传递硬件上执行的处理的部分。在图 3 所示的实施例中,在块 330 将命令放置在传递队列中之后,在块 340 控制器确定是否使能中断。下面参照图 4 进一步详细讨论是否使能中断的决定。如以下进一步描述的,使能或者禁止中断指控制器是否对传递硬件产生的中断做出响应。在本实施例中,控制器 100 是否对传递硬件 110 产生的中断做出响应是通过屏蔽或者不屏蔽来自传递硬件 110 的(若干)中断比特控制的。

[0021] 如果在块 340 控制器确定中断是必要的,则在块 350 控制器不屏蔽(若干)中断比特,在块 370 控制器针对屏蔽状态检查设置“未屏蔽的”状态。如果在块 340 确定不需要中断,则在块 360 屏蔽(若干)中断比特,在块 370 针对屏蔽状态检查设置“屏蔽的”状态。如果(若干)中断比特被屏蔽,则当传递硬件产生中断时控制器继续其它任务并且不被中断。例如,在块 360 屏蔽(若干)中断比特之后,在块 400 控制器可以接着检查分级队列中是否存在任何命令。在另一个实施例中,可根据不同的逻辑屏蔽/不屏蔽比特以指示使能/禁止。在其它进一步实施例中,不屏蔽比特,替代地控制器可以使用其它机制(例如,在硬件或者存储器中设定特定值)以指示使能或者禁止中断。

[0022] 在这个实施例中,当在块 420 传递硬件 110 执行主机传递时,如在图 2 中那样,传递硬件 110 继续进行以产生中断 430。在控制器对中断做出响应之前,在块 370 中断现在通过屏蔽状态检查。如果在屏蔽状态检查 370 确定与中断相关联的(若干)中断比特已经被屏蔽,则控制器不进行进一步动作(如块 410 所示)并且控制器继续执行其它处理。如果在块 370 屏蔽状态检查确定所产生的中断与未屏蔽的(若干)中断比特相关联,则中断被控制器处理。在一个实施例中在块 380 控制器接着产生上下文切换并且进行到块 390 和 400,如以上关于图 2 描述的。

[0023] 在其他实施例中,可使用多个替代配置以减少控制器 100 处理的硬件中断的数量。尽管图 3 的实施例示出中断总由传递硬件产生并且被控制器选择性地处理,但是其它实施例可以允许控制器停止在传递硬件侧产生中断。在一些实现方式中,可以基于单独的主机传递使能或者禁止硬件中断,但是在其它实现方式中,可以基于一批主机传递使能或者禁止硬件中断。本领域技术人员将理解可根据此处描述的总体配置使用选择性地使能硬件中断的其它方法以减少主机传递完成时控制器的中断。

[0024] 在一个实施例中,可能在图 3 所示的处理命令之前必须针对每个标识标签建立资源。这例如可通过在装置加电时针对每个标识标签分配特定存储器资源(例如,在操作传递存储器的缓冲器中)完成。例如,每个标识标签可以被提供有缓冲存储器的专用部分。作为替代,控制器可以被配置以在实现和不实现选择性中断方案的操作模式之间交替。在总中断方案已经为每个标识标签分配资源之后,可激活选择性中断方案。

#### [0025] 禁止主机传递中断的条件

[0026] 图 4 是示出描绘根据实施例的基于特定条件确定是否使能主机传递中断的处理的流程图。该处理例如可作为图 3 的块 340 中进行的确定的一部分执行。如图 4 的实施例中所示,在块 500 新命令被处理接收,该处理确定是否使能主机传递中断。在块 510,确定主机传递是否和读命令相关。在一个实施例中,对于读命令,主机传递中断可能是不必要的,因为在执行读命令主机传递时,先前从介质读取的数据被传递到主机。在完成传递时,控制器不要求进一步的动作来完成读命令。相反地,对于写命令,在特定实施例中,主机传递从主机收集主机已经请求的数据,以写入到介质存储器。因此在接收数据之后控制器要求额外的动作以实现数据的存储。主机传递完成中断可触发向存储介质的写命令数据传递。因此,如果命令不是读命令,则在块 540 主机传递中断被使能。如果主机传递命令是读命令,则处理进行到块 520。

[0027] 在块 520,一个实施例中的处理确定该读命令是否是过大的读命令。在示例实施例中,比可在单个主机传递请求中传递的数据量大的读命令被认为是过大的读命令。对于这种过大的读命令,在主机传递完成时控制器可能需要被中断,以产生包括剩余读数据的另一个主机传递。因此,对于过大的读命令,在块 540 主机传递中断可被使能。对于正常大小的读命令,处理继续到块 530。

[0028] 在块 530,处理在一个实施例中确定这个命令的添加是否将填充主机传递队列。如果主机传递队列将是满的,则在添加当前命令之后不可向队列添加额外的传递。因此,额外的传递可以存储在控制器上在分级队列中等待主机传递队列中的时隙。可以在此情形中产生中断以用信号向控制器通知在主机传递完成时时隙变得可用。根据此实施例,当在添加进入的命令之后主机传递队列将变为满时,在块 540 主机传递中断可以被使能。当通过添加进入的命令主机传递队列将不变为满时,主机传递中断将被禁止,如块 550 指示的。

[0029] 根据利用控制器和接口的实现方式细节,在本发明的其它实施例中可以改变图 4 中所示的条件。作为示例,特定数据传递接口可能不允许使用过大的读命令,在此情况下检查过大的命令将不是必要的并且如果满足特定的其它条件则控制器将禁止针对全部读命令的中断。在其它实现方式中,如果写数据直接被传到介质而无需任何必要的控制器动作,则写命令主机传递可能不需要中断(因而被禁止)。作为进一步示例,主机传递队列可以具有足够的深度从而分级队列是不必要的,或者否则不能够变为不能接受额外请求。在该情

况下,不需要块 530 中进行的检查。因此,在各种实施例中,在图 4 的确定处理中进行的检查的类型可以不同和 / 或按照与所描绘的顺序不同的顺序进行。

[0030] 结论

[0031] 以上公开的特定实施例的特征和属性可以按照不同方式组合以形成附加实施例,全部实施例落入本公开的范围。尽管本公开提供特定实施例和应用,但是其它实施例对于本领域技术人员是明显的,包括不提供此处阐述的全部特征和优点的实施例,并且它们也在本公开的范围。因此,本公开的范围旨在仅参照所附权利要求限定。



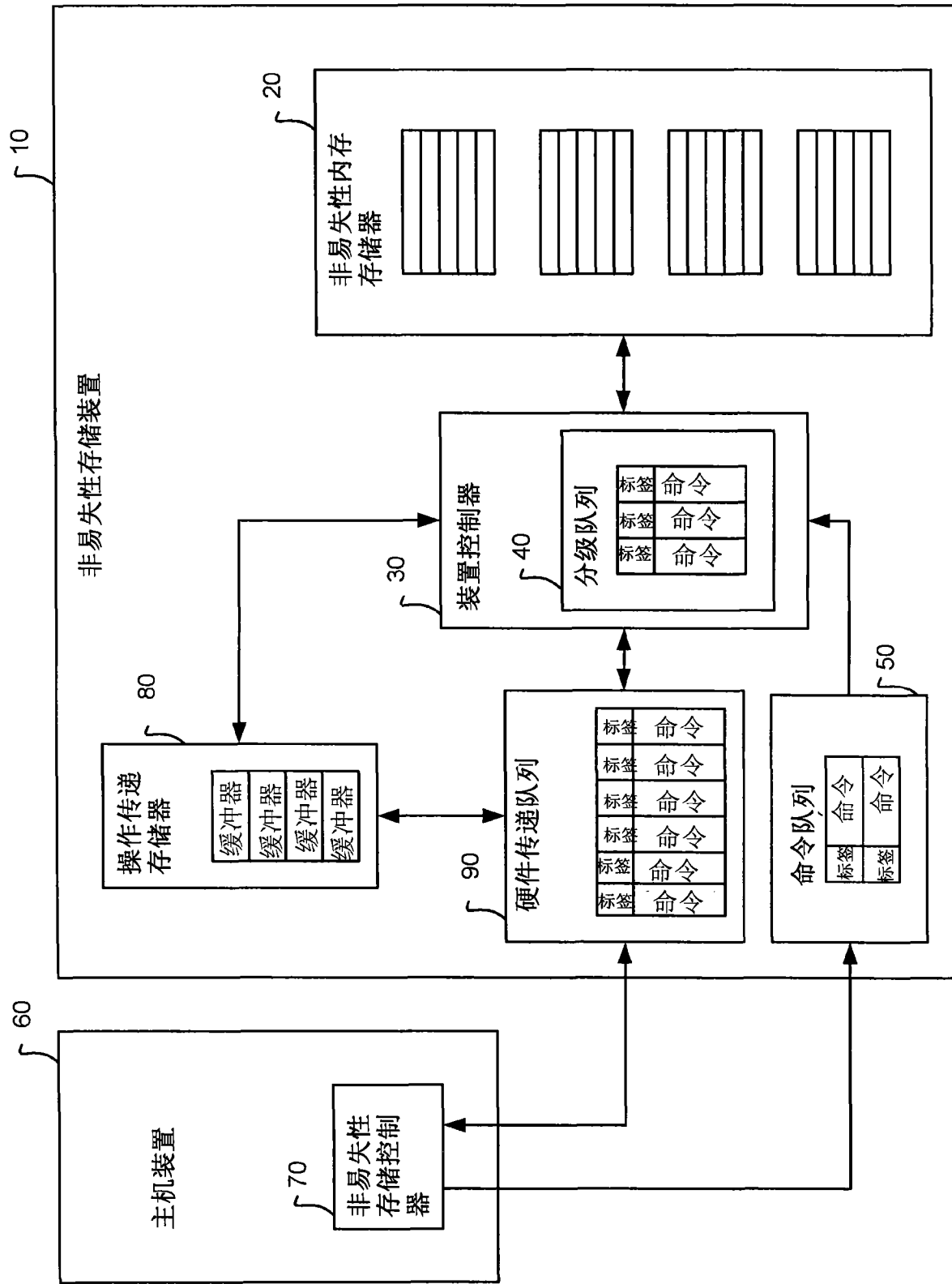


图 1

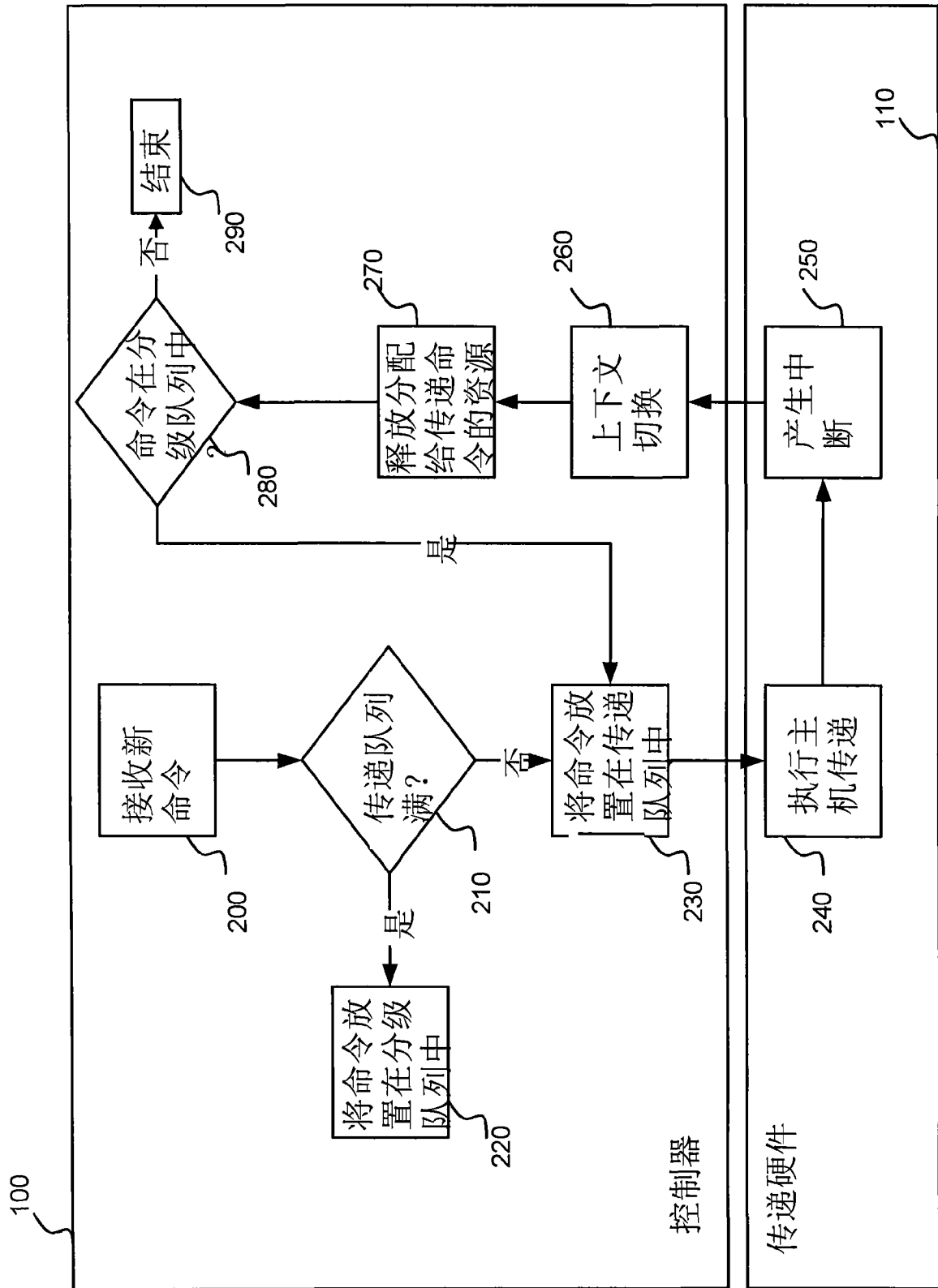


图 2

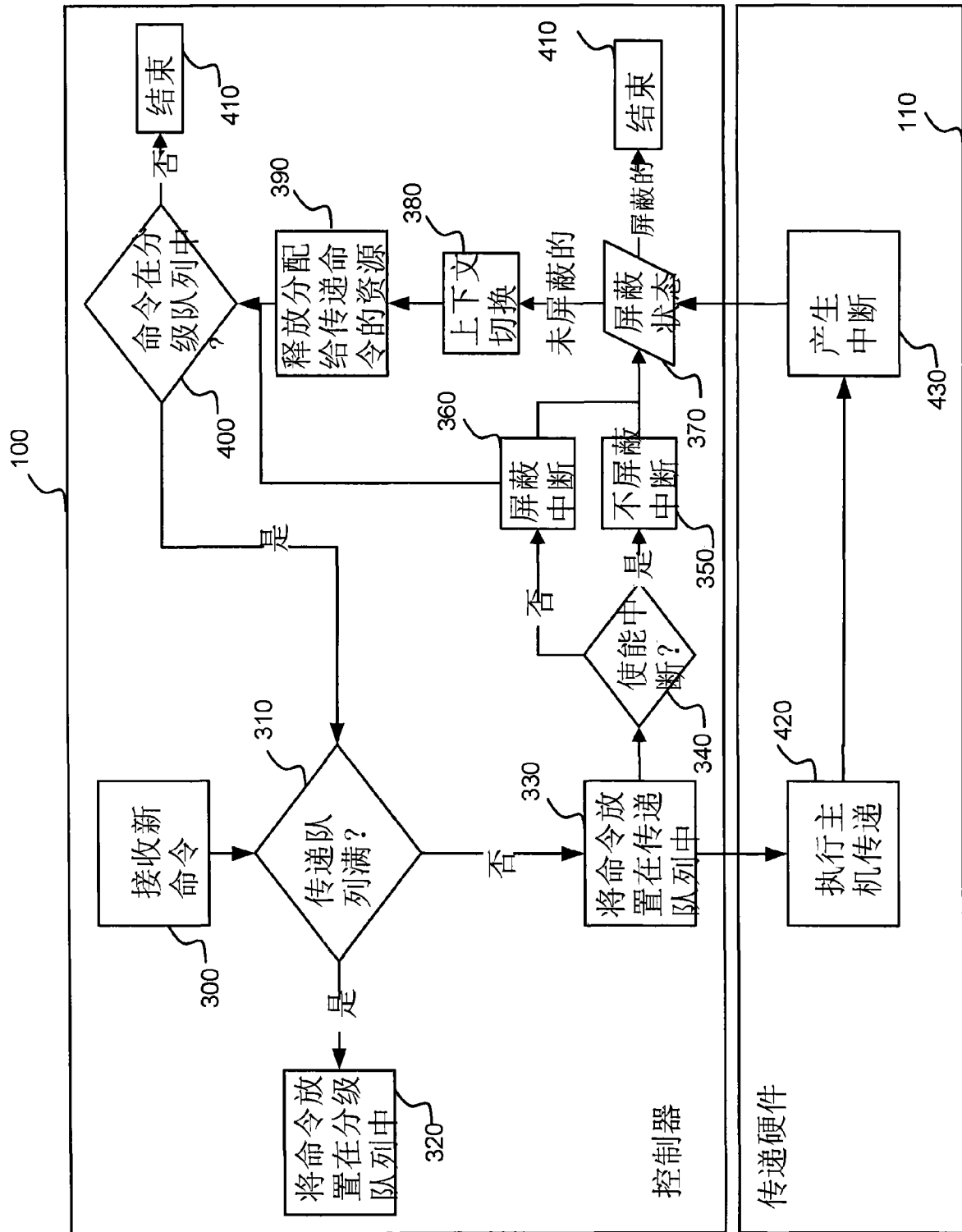


图 3

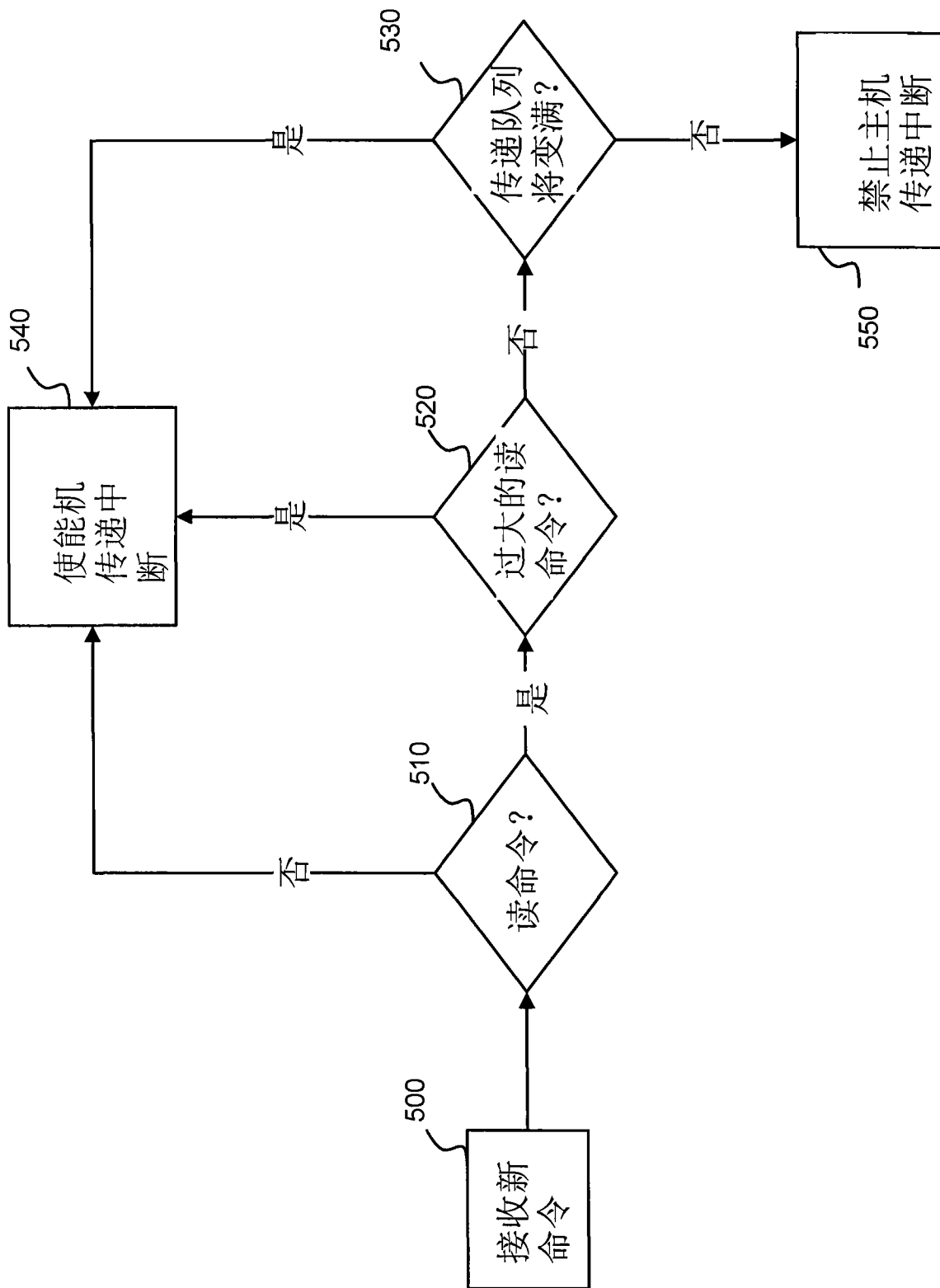


图 4