



(12)发明专利

(10)授权公告号 CN 104008073 B

(45)授权公告日 2017. 12. 12

(21)申请号 201410058457.8

(22)申请日 2014.02.20

(65)同一申请的已公布的文献号

申请公布号 CN 104008073 A

(43)申请公布日 2014.08.27

(30)优先权数据

13/773,337 2013.02.21 US

(73)专利权人 希捷科技有限公司

地址 美国加利福尼亚州

(72)发明人 M·R·弗伦杰尔姆

(74)专利代理机构 上海专利商标事务所有限公

司 31100

代理人 何焜

(51)Int.Cl.

G06F 13/16(2006.01)

G06F 9/46(2006.01)

G06F 9/445(2006.01)

(56)对比文件

US 2012/0054739 A1,2012.03.01,

US 2010/0199268 A1,2010.08.05,

CN 101305334 A,2008.11.12,

US 2008/0243947 A1,2008.10.02,

US 2008/0162817 A1,2008.07.03,

审查员 贾东曜

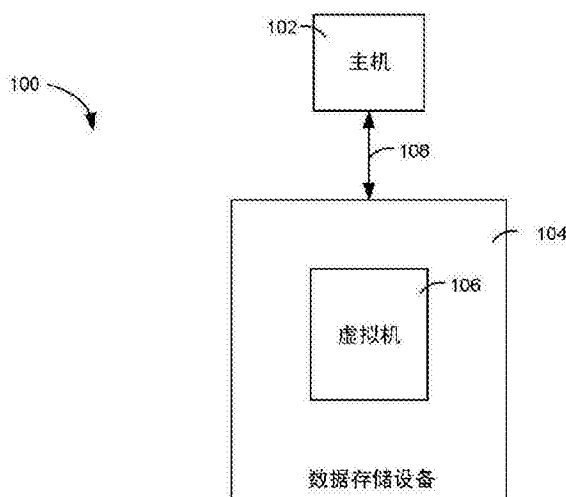
权利要求书3页 说明书5页 附图3页

(54)发明名称

具有虚拟机的CDATA存储设备

(57)摘要

本申请公开具有虚拟机的CDATA存储设备。公开了在数据存储设备上实施虚拟机的装置、系统和方法。在一个实施例中,设备可以包括响应于主机的通信接口、处理器和壳体,所述壳体包括通信接口和处理器,使得该设备可从主机移除。所述处理器可经配置以通过通信接口从主机接收命令,使用不依赖于设备架构的独立平台的编程接口处理命令,并通过所述通信接口向主机返回命令结果。在另一实施例中,一种方法可包括:在数据存储设备从主机设备接收命令,在数据存储设备使用独立于平台的编程接口处理所述命令,以及从所述数据存储设备向所述主机返回命令结果。



1. 一种数据存储驱动器,包括:

响应于主机的通信接口,且所述通信接口配置为接收与用户数据相关联的命令,所述命令指示要由虚拟机执行的代码;

虚拟机,该虚拟机包括所述数据存储驱动器的独立于平台的编程接口,所述编程接口不依赖于所述数据存储驱动器的架构;

处理器,经配置以:

通过所述虚拟机接收所述命令以产生计算结果,该计算结果包括新产生的数据,该数据不是写入命令响应且不是预先存储的数据;

通过所述通信接口向所述主机提供所述命令的所述计算结果;以及

壳体,包括所述通信接口和所述处理器,使得所述数据存储驱动器从主机是可移除的。

2. 如权利要求1所述的数据存储驱动器,进一步包括连接到所述壳体的数据存储介质。

3. 如权利要求2所述的数据存储驱动器,进一步包括:所述处理器经配置以:

确定该命令是应用程序命令或数据存储命令,所述应用程序命令包括将被虚拟机执行以产生所述计算结果的操作,以及所述数据存储命令包括读命令或写命令;

当所述命令是应用程序命令时,使用所述虚拟机处理命令;以及

当所述命令是数据存储命令时,向所述数据存储驱动器存储或从所述数据存储驱动器检索数据。

4. 如权利要求2所述的数据存储驱动器,进一步包括:通过所述虚拟机处理命令包括:

从数据存储介质检索与所述命令相关联的所述用户数据,以及

根据所述命令处理所述用户数据。

5. 如权利要求1所述的数据存储驱动器,进一步包括:所述命令包括计算机代码以及利用所述虚拟机执行所包含的计算机代码的指令。

6. 如权利要求1所述的数据存储驱动器,进一步包括:所述虚拟机包括被配置成执行指令集的逻辑电路。

7. 如权利要求1所述的数据存储驱动器,进一步包括:所述虚拟机包括在处理器上运行的计算机可读指令。

8. 如权利要求1所述的数据存储驱动器,进一步包括:

所述命令包括要被处理的数据以及如何处理数据的指令,以及

所述处理器进一步经配置以根据所述指令使用所述虚拟机来处理所述命令。

9. 如权利要求1所述的数据存储驱动器,进一步包括:所述命令包括指令以在所述主机上执行过程的子过程以分配计算工作量。

10. 如权利要求1所述的数据存储驱动器,进一步包括:所述虚拟机不依赖于所述数据存储驱动器的操作系统。

11. 一种计算机系统,包括:

主机;

数据存储驱动器,配置为在其上存储用户数据,包括:

响应于所述主机的通信接口,且所述通信接口配置为接收与用户数据相关联的命令,所述命令指示要由虚拟机执行的代码;

虚拟机,该虚拟机包括所述数据存储驱动器的独立于平台的编程接口,所述编程接口

不依赖于所述数据存储驱动器的架构；

处理器,经配置成:

通过所述虚拟机接收所述命令以产生计算结果,该计算结果包括新产生的数据,该数据不是写入命令响应且不是预先存储的数据;

通过所述通信接口向所述主机提供所述命令的所述计算结果;以及

包括接口和处理器的壳体,使得所述数据存储驱动器是从主机可移除的。

12. 如权利要求11所述的系统,所述数据存储驱动器进一步包括:连接到所述壳体的数据存储介质。

13. 如权利要求12所述的系统,进一步包括:所述处理器进一步被配置以:

确定该命令是应用程序命令或数据存储命令,所述应用程序命令包括指定要由虚拟机执行的操作的命令,以及所述数据存储命令包括读命令或写命令;

当所述命令是应用程序命令时,使用虚拟机处理所述命令;以及

当所述命令是数据存储命令时,向数据存储介质存储数据或从数据存储介质检索数据。

14. 如权利要求12所述的系统,进一步包括:使用所述虚拟机处理命令包含:

从所述数据存储介质检索与所述命令相关联的数据,以及

根据所述命令处理数据。

15. 如权利要求11所述的系统,进一步包括:所述主机经配置以:

向所述数据存储驱动器存储数据,以及

发出命令到所述数据存储驱动器,所述命令包括使用所述虚拟机处理原始数据的指令。

16. 如权利要求11所述的系统,进一步包括:所述主机经配置以:

在所述主机处执行包括多个子过程的过程;以及

发出命令到数据存储驱动器,所述命令包括在所述虚拟机上执行多个子过程之一的指令。

17. 如权利要求11所述的系统,进一步包括:所述虚拟机不依赖于数据存储驱动器的操作系统。

18. 一种用于在数据存储驱动器上实施的方法,包括:

在数据存储驱动器从主机设备接收命令,所述数据存储驱动器配置为在其上存储用户数据,且可从所述主机设备移除所述数据存储驱动器;

在所述数据存储驱动器,使用不依赖于数据存储驱动器的架构的独立于平台的编程接口处理所述命令,所述处理产生包括新数据的结果,该新数据不是写入命令响应,且不是预先存储的数据,以及

从数据存储驱动器向主机返回命令结果。

19. 如权利要求18所述的方法,进一步包括:在数据存储驱动器使用独立于平台的编程接口处理命令的方法,包括:

从所述数据存储驱动器检索命令相关联的数据,以及

根据所述命令处理数据。

20. 如权利要求18所述的方法,进一步包括:所述命令包括用于在数据存储驱动器以执

行在主机上运行的过程的子过程的指令,以分配计算工作量。

具有虚拟机的CDATA存储设备

背景技术

[0001] 本发明一般涉及数据存储设备上的虚拟机和独立于设备的编程接口。

发明内容

[0002] 一般地,提出了一种包括虚拟机的装置数据存储设备,所述虚拟机经配置以允许独立于设备的编程接口。在一个示例中,设备可以包括响应于主机的通信接口,处理器和壳体,所述壳体包括通信接口和处理器,使得该设备可从主机移除。所述处理器可经配置以通过通信接口从主机接收命令,使用不依赖于设备架构的独立平台的编程接口处理命令,并通过所述通信接口向主机返回命令结果。

[0003] 在另一示例中,系统可以包括主机和数据存储设备,所述数据存储设备包括:响应主机的通信接口,处理器以及外壳,所述壳体包括接口和处理器,使得数据存储设备可从主机移除。所述处理器可经配置以通过通信接口从主机接收命令,使用不依赖于数据存储设备的结构架构的虚拟机处理命令,并通过所述通信接口向主机返回命令结果。

[0004] 在又一示例中,一种方法可包括:在数据存储设备从主机设备接收命令,在数据存储设备使用不依赖于数据存储设备的架构的独立平台的编程接口处理所述命令,以及从所述数据存储设备向所述主机返回命令结果。

附图说明

[0005] 图1是具有虚拟机的数据存储设备的示例性实施例的示意图;

[0006] 图2是具有虚拟机的数据存储设备的另一示例性实施例的示意图;以及

[0007] 图3是用于实施具有虚拟机的数据存储设备的方法的示例性实施例的流程图。

具体实施方式

[0008] 在本实施例的以下详细描述中,参考了形成本文一部分并且其中通过具体实施例的图示进行说明的附图。但是应当理解的是:在不脱离本发明的精神和范围的情况下,可以利用和改变其他实施例。

[0009] 参照图1,计算机系统的实施例示出并且一般指定为100。系统100可以包括主机102,数据存储设备(DSD)104以及接口108。主机102也可以被称为主机系统或主机计算机。主机102可以是台式计算机、膝上型计算机、服务器、个人数字助理(PDA)、电话、音乐播放器、另一个电子设备或它们的任意组合。DSD104可以是磁盘驱动器、固态驱动器、混合动力驱动、USB闪存驱动器或任何其他类型的存储设备。该DSD104可以是与主机102单独可移动的,并且可以包括外壳,壳体或包围DSD104的组件的其他物理外壳。接口108可以是主机102和DSD104之间的通信介质,并且可以包括允许DSD104连接于或脱离于主机102的连接器的接口204可以包括允许主机202和DSD200之间通信的有线或无线的任何接口,诸如USB、IEEE1394、压缩型闪存、SATA、eSATA、PATA、SCSI、SAS、PCIe、光纤通道、以太网或thunderbolt等等。

[0010] 所述数据存储设备104可以包括虚拟机编程接口(VM)106。虚拟机可是独立于平台的编程环境,该编程环境提供了可以掩盖底层硬件或操作系统(OS)的细节并允许为虚拟机而不是为之上VM正运行的底层计算机平台设计程序的抽象等级。平台可以指设备的硬件结构和软件框架,诸如计算机的物理组件和操作系统。通过独立于平台,无论系统的架构或操作系统,虚拟机环境可以被设计以呈现相同的编程接口。换句话说,计算机应用可以被设计成与特定虚拟机接口并在其上执行,而不要求其将运行的系统的细节。

[0011] 在一些实施例中,虚拟机可以模仿其他设备的平台特性,即使运行VM106的实际设备不共享该特性。也就是说,设计成操作特定平台的应用可以在在模仿该平台的虚拟机上运行,VM运行在具有不同平台的设备上。这仍然是独立于平台的,这是因为设备通过使用VM106仍可运行为不同平台设计的应用。

[0012] 在一些实施例中,虚拟机可以是提供程序环境以运行处理的“过程”虚拟机,或是模拟或模糊硬件架构的“系统”虚拟机,诸如硬件抽象层(HAL)。

[0013] VM106可以是任何类型的虚拟机环境,包括Java虚拟机、Dalvik、运行公共语言运行库的微软的.net框架以及Parrot虚拟机,等等。应用程序可以被设计为运行在VM环境106中,而不需要重新设计应用以考虑DSD104的特定物理组件(诸如处理器)、操作系统等。这可以允许使用DSD作为应用的处理设备,而不需要根据设备制造商或架构重新设计应用程序。

[0014] 在一个示例中,有两个设备:一个是运行Windows8操作系统的移动电话,而另一个是运行Linux操作系统的计算机服务器,并且两个设备都具有Java虚拟机。尽管架构和操作系统的差异,经设计以在Java VM上运行的应用程序可以运行在两台设备上,而不需要应用程序本身的任何更改。

[0015] 虚拟机106可以被实施以在处理器、控制器、数字信号处理器、精简指令集计算机(RISC)、专用集成电路(ASIC)、现场可编程门阵列(FPGA),系统单一芯片、或者具有足够计算能力的任何引擎上运行。VM106可以被实现为在DSD104上运行的软件,或者作为具体配置的电路。在一些实施例中,DSD104可以运行操作系统,VM106运行在操作系统环境中并同时提供独立于OS的编程接口。

[0016] 参照图2,具有虚拟机的数据存储设备的特定实施例并示出并且一般指定为200。具体而言,图2提供了盘驱动器的数据存储设备(DSD)200的功能框图。该DSD200可以是数据存储设备,诸如图1中所示的设备104。数据存储设备200可以通过基于硬件或固件的接口电路204与主机设备202(诸如在图1中示出的主机系统102)进行通信,所述接口电路204可包括允许DSD200物理上从主机202去除的连接器(未示出)。在一些实施例中,DSD200可以通过有线或无线通信经过接口204与主机202进行通信。缓冲器212在读写操作或虚拟机处理期间临时存储数据,并且可以包含命令队列(CQ)213,其中多个挂起操作可以暂时储存待处理执行。非易失性固态存储器203(诸如闪存)可以被包含用于额外的缓存或缓冲存储器,或为DSD200提供额外的数据存储。包含多种类型的非易失性存储介质(诸如盘209和闪存203)的DSD200可以是混合的存储设备。

[0017] 该DSD200可以包括具有相关的存储器208和处理器210的可编程控制器206。控制器206可以实施虚拟机环境211为独立于平台的编程接口。在一些实施例中,DSD200可包括处理器205,以及虚拟机211可以被实施为在处理器205上运行的计算机指令,或作为FPGA、ASIC等。

[0018] 此外,图2示出:DSD200可以包括读写(R/W)信道217,它可以在写操作期间编码读取并在读操作期间重建从盘209中检索的用户数据209。前置放大器电路(前置)218可以向头部219施加写入电流,并提供前置放大回读信号。伺服控制电路220可以使用伺服数据以向电流线圈224提供合适电流用于定位头部219。在执行命令队列213中的各个待处理命令期间,控制器206可以与处理器222进行通信以将头部219移动到盘209上的所需位置。

[0019] 在示例实施例中,在DSD200从主机202通过接口204接收命令。接口204可以包括允许主机202和DSD200之间通信的有线或无线的任何接口,诸如USB、IEEE1394、压缩型闪存、SATA、eSATA、PATA、SCSI、SAS、PCIe、光纤通道,以太网或Thunderbolt等等。该DSD200可经配置以确定命令的性质和相应行为。例如,命令可以是要求在DSD200存储或检索数据的数据存储命令。该命令还可以是可以在虚拟机211上运行的任务或应用程序,并且还可以包括与该任务或应用程序相关联的数据。在一些实施例中,数据存储命令可包括在通信接口的标准指令集中包括的指令,而VM任务可以是非接口标准的命令。

[0020] 例如,来自主机的样本VM命令可以由VM执行的代码或在设备上存储将被VM执行的代码的指针组成。例如,VM命令可以包含Java字节代码的数据,当由DSD接收时直接由DSD中的Java VM解释。主机还可首先发出“标准”的写数据命令以将数据写入到存储设备上的一个或多个块,并然后发出VM命令,指示DSD从存储设备加载一个或多个块并执行那里存储的代码。此外,因为VM可以在存储设备上实施存储空间,主机命令可以通过VM文件名而不是块号指向数据。

[0021] 在使用标准接口(诸如SATA或SAS)的VM的示例实施方式中,VM相关命令的执行可仅仅是命令集的扩展。例如,在SAS中使用的SCSI命令集是由T10委员会定义的,并且包括(对于诸如DSD的块设备)定义命令用于读取和写入块的SBC-3标准。该命令可以传送到使用CDB(命令描述块)的设备。例如,SCSI命令集的“读取”命令是6个字节的序列,其中第一字节是命令代码(用于读取,该字节的值是“6”),第二、第三和第四字节是逻辑块地址,第五字节是传输长度以及最后一个字节是“控制”。以类似的方式,在SCSI命令集中定义“写”、“格式单位”以及其他命令。因此,实施VM命令的示例方法将使用命令字节的新值定义新的命令(有256个可能的字节),并定义了命令如何相应地工作。

[0022] 在示例实施例中,可以使用SCSI定义VM。当主机连接到SCSI设备时,它可使用SCSI查询命令,它有自己的12值以识别所连接的设备。如果它是直接存取块设备,该设备然后用00响应。有直接存取块设备类型的多个子类,以及可以加入新类型以定义“直接访问块设备与虚拟机能力”类型,该类型可以在设备中实施VM操作所需要的一组命令。在CDB中主机为命令字节使用特定值的事实将使得存储设备识别命令的类型。

[0023] 另一个示例实施方式是通过以太网接口和TCP/IP协议,其中主机通过发现机制(诸如广播,或已知VM设备的主机定义列表)会知道实施VM的设备的IP地址,并然后能直接将数据包发送给连接到设备内VM的TCP/IP端口。

[0024] 如果该命令是读或写命令,控制器206可指示在盘209、闪存203、或其他存储器的相关数据的检索或存储。在完成后,DSD200可以返回所请求的数据或返回的数据已经被存储的通知。

[0025] 如果命令包含任务用于在VM211上实施,VM211的实例可以被创建,或者任务可被发送到已经运行的VM211的实例。VM211可以在控制器206、单独的处理器205或其它计算组

件上实施。任务或应用程序可以采用将由VM211执行的指令或实施VM211的给定功能的指令的形式。该命令除了任务可以包括一组数据,以及该任务可对数据执行操作。例如,任务可以是等式以对所提供的数据集执行。在一些实施例中,该命令可指示该VM211例如通过提供一系列的逻辑块地址(LBA的)而对DSD200上存储的数据执行操作。

[0026] 在一个实施例中,大的数据集可被存储在DSD200上,以及任务可以是比较所存储的数据集与由主机提供的数据集,并返回最接近的匹配。在另一个实施例中,任务可以是等式以对每个数据集执行并把结果返回给主机。另一个实施例可涉及根据由主机202提供的标准而排序所存储的数据。在一些实施例中,该任务可涉及:将结果存储在DSD200上,并向主机202返回所存储结果的地址,而不是向主机返回结果本身。

[0027] 返回给主机的结果取决于从主机接收的命令是读、写命令或者VM任务而有所不同。例如,在读命令后,DSD200可以返回所请求的存储数据,以及在写命令后,DSD200可返回写入完成响应。在该命令是VM任务的实施例中,返回的结果可是在VM211上执行的任务或计算的结果,或任务或计算的结果的存储位置,或其他非接口标准的结果。例如,响应于VM命令返回的数据可不包括在接收命令之前写入完成响应或DSD200中存储的数据,而是包括响应于所述命令新产生的数据。

[0028] 在VM211内执行任务或算法可以使得在DSD200快速处理,而不需要通过接口204向主机发送大量数据。因此,可以利用靠近大批量数据执行的任何操作都可受益于在DSD上实施的VM,诸如对较大数据集执行的搜索功能,其中在主机只需要积极匹配。

[0029] 在一些实施例中,发送到DSD200的任务包括由主机202所执行的过程的子过程。例如,在主机202上运行的程序可包括几个进程线程,其子集可以发送到DSD200进行处理。在另一实施例中,主机202可连接到具有虚拟机三维多个数据存储设备,或具有VM的多个DSD可以其他方式连接在网络中。每个DSD上的VM211可经指示以执行较大过程的子过程,所述较大过程需要多次计算或大批量数据。例如,必须处理的大量数据可以存储在多个数据存储设备,以及每个DSD被指示以处理其相应的存储的数据。该系统将通过分配计算工作量、通过由在其中存储数据的驱动上执行处理(而不是通过数据总线或可阻塞理的接口发送大量数据)而减少延迟来改进数据处理时间。

[0030] 例如,数据库可包含由服务器管理的1,000TB的数据,所述服务器连接到每个具有1TB的1000个驱动器。如果搜索是要由服务器执行,它可需要从一个驱动器一次向服务器顺序发送(例如,以100MB/s)1,000TB的数据。以这种方式进行搜索可需要数天才能完成。如果每个驱动器具有能对其各自的数据执行搜索功能的VM,搜索可以在数小时内完成。

[0031] 另外,功能可以比搜索更加复杂。例如,连锁店可以有销售记录的数据库,并提供分布在具有VM的存储设备之间的信息,该信息可以被访问以根据供应、需求和以往的销售历史价格提供最新的价格预测。如果数据必须通过有限带宽传输并通过中央服务器进行处理,该计算可能过于缓慢。涉及存储在多个驱动的大量数据的“大数据”的类似其他应用也是可能的,例如,使用互联网访问历史、用户偏好及其对等网络来选择合适的广告,朋友推荐,音乐选择等。

[0032] 虚拟机也可以实施NAS-网络附加存储设备-其中提供存储用于其他设备进行备份、媒体存储、精简配置(thin provisioning)、重复数据删除和冗余,该方式具有额外的好处:代码实施跨设备供应商和未来技术是可传输的。

[0033] 参考图3,示出了用于在数据存储设备上实施虚拟机的方法300的示例性实施例的流程图。该方法300可以包括:在302,从主机接收命令。该方法300可以包括:在304,确定该命令是用于由数据存储设备上运行的虚拟机实施的应用程序指令,或诸如数据读或写命令的数据存储命令。如果在306确定该命令是数据存储命令,该方法可涉及:在308,向数据存储设备存储相关联的数据或从数据存储设备中检索数据。

[0034] 如果在306确定该命令不是数据存储命令,在310,该命令可以在VM上执行的应用程序指令。应用程序指令可采用计算机可执行代码、指令、或指定要在VM上执行的任务的其他指示的形式。例如,应用程序指令可包括一组将要对数据集进行的数学函数。应用程序指令还可以包括在其上执行任务或操作的数据,或者它可以表示其中在数据存储设备上存储该数据的地址。应用程序指令也可指示任务的结果是否要被返回到主机、存储到数据存储设备或两者。如果结果要被存储到数据存储设备上,在312,返回该任务完成或未能完成的通知或存储结果的地址。

[0035] 在308完成数据存储命令或在310的应用程序命令之后,在312向主机返回相应的结果或数据。

[0036] 根据各种实施例,本文描述的方法可以被实现为运行在计算机处理器或控制器(诸如控制器206)的一个或多个软件程序。专用硬件实现方式包括(但不限于)专用集成电路,可编程逻辑阵列和其他硬件设备同样可以构造成实现本文所描述的方法。进一步,本文描述的方法可以被实现为包括指令的计算机可读介质,当执行所述指令时引起处理器执行该方法。

[0037] 本文所述实施例的说明旨在提供对各种实施例的结构的一般理解。本说明并非旨在作为利用本文描述的结构或方法的装置和系统的所有元件和特征的完整描述。本领域技术人员在浏览本公开之后,许多其他实施例可以是显而易见的。其他实施例可以被利用并源自本公开内容,使得在不脱离本发明的精神和范围下可以进行结构和逻辑替换和改变。而且,尽管具体实施例已被图示和描述,但应当理解,设计以实现相同或类似目的的任何随后配置可替代所示的具体实施例。

[0038] 本公开内容旨在涵盖任何及所有的后续的修改或各种实施例的改变。在本领域技术人员浏览本描述之后,上述实施例的组合以及本文中未具体描述的其它实施例将是显而易见的。此外,该图示仅是代表性的,并且可以不按比例绘制。图内的某些比例可被放大,而其他比例可减少。因此,本公开和附图应被视为说明性的而不是限制性的。

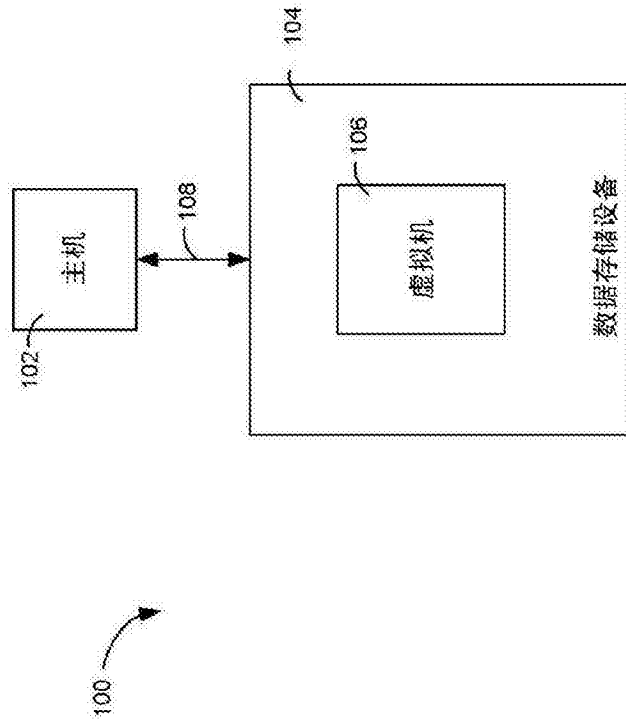


图1

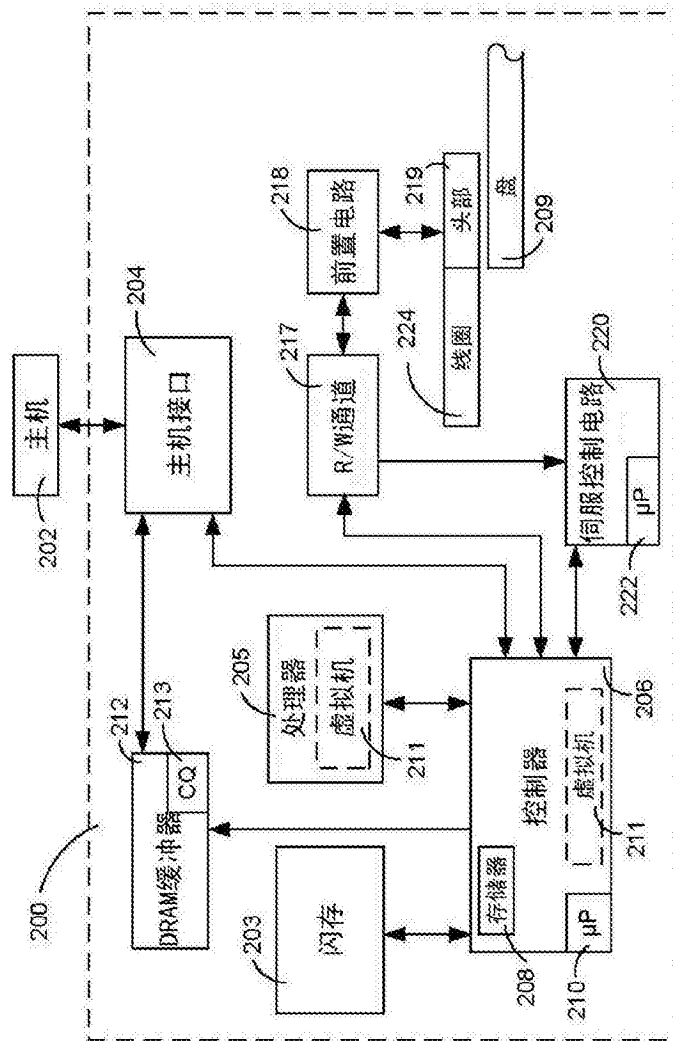


图2

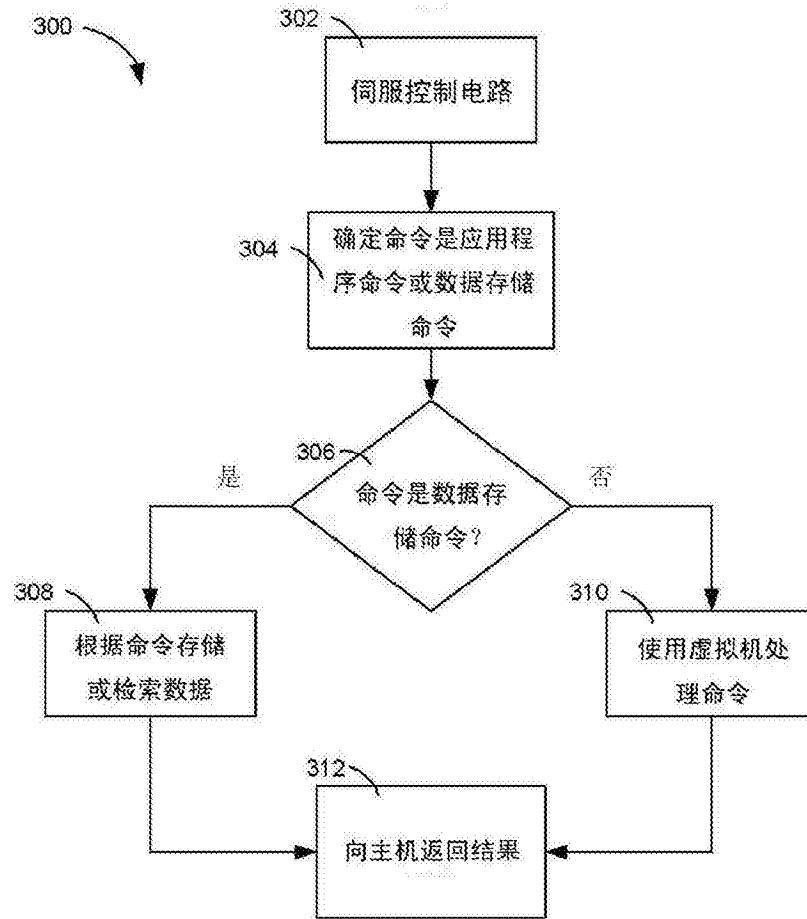


图3