

一种云环境下的 I/O 卸载方法、设备、系统及存储介质

本申请要求于 2022 年 08 月 30 日提交中国专利局、申请号为 202211060455.3、申请名称为“一种云环境下的 I/O 卸载方法、设备、系统及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

本申请涉及云计算技术领域，尤其涉及一种云环境下的 I/O 卸载方法、设备、系统及存储介质。

背景技术

随着 CPU 的性能增长边际成本急剧上升，I/O 卸载卡应用而生。I/O 卸载卡，又称 DPU、IPU，可为高带宽、低延迟、数据密集的计算场景提供 I/O 计算引擎。

目前，I/O 的处理工作完全卸载到了 I/O 卸载卡中，I/O 卸载卡上的 I/O 处理工作与宿主机的 CPU 数据链路分离，这种分离结构导致目前的 I/O 处理过程中经常发生数据溢出等问题，处理效率不佳。

发明内容

本申请的多个方面提供一种云环境下的 I/O 卸载方法、设备、系统及存储介质，用以改善云环境下的 I/O 处理效率。

本申请实施例提供一种云环境下的 I/O 卸载系统，包括：宿主机的 CPU 以及插接在所述宿主机上的 I/O 卸载卡，所述 CPU 内装配有队列组件；

所述队列组件，用于为所述宿主机上的实例提供虚拟 I/O 设备以及与所述虚拟 I/O 设备对应的设备队列；在所述 CPU 内，利用所述设备队列对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求进行调度；

所述 I/O 卸载卡，用于监听所述设备队列中的 I/O 请求；在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

本申请实施例还提供一种云环境下的 I/O 卸载方法，适用于宿主机中的 CPU，所述 CPU 中装配有队列组件，所述方法包括：

- 利用所述队列组件为所述宿主机上的实例提供虚拟 I/O 设备；
- 利用所述队列组件为所述虚拟 I/O 设备配置对应的设备队列；
- 在所述 CPU 内，利用所述设备队列对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求进行调度，以供所述宿主机上插接的 I/O 卸载卡监听所述设备队列中的 I/O 请求

并在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

本申请实施例还提供一种云环境下的 I/O 卸载方法，适用于宿主机上插接的 I/O 卸载卡，所述宿主机的 CPU 中装配有队列组件，所述方法，包括：

- 5 从所述队列组件中的设备队列中监听所述宿主机上的实例与所述队列组件为所述实例提供的虚拟 I/O 设备之间发生的 I/O 请求；

获取监听到的 I/O 请求所相关的数据；

在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

- 10 本申请实施例还提供一种处理器 CPU，安装在宿主机中，所述 CPU 内装配有队列组件，所述 CPU 用于执行所述一条或多条计算机指令，以用于：

利用所述队列组件为所述宿主机上的实例提供虚拟 I/O 设备；

利用所述队列组件为所述虚拟 I/O 设备配置对应的设备队列；

- 15 在所述 CPU 内，利用所述设备队列对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求进行调度，以供所述宿主机上插接的 I/O 卸载卡监听所述设备队列中的 I/O 请求并在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

本申请实施例还提供一种 I/O 卸载卡，插接于宿主机上，所述宿主机的 CPU 内装配有队列组件，所述 I/O 卸载卡包括存储器和处理器；

- 20 所述存储器用于存储一条或多条计算机指令；

所述处理器与所述存储器耦合，用于执行所述一条或多条计算机指令，以用于：

从所述队列组件中的设备队列中监听所述宿主机上的实例与所述队列组件为所述实例提供的虚拟 I/O 设备之间发生的 I/O 请求；

获取监听到的 I/O 请求所相关的数据；

- 25 在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

本申请实施例还提供一种存储计算机指令的计算机可读存储介质，当所述计算机指令被一个或多个处理器执行时，致使所述一个或多个处理器执行前述的云环境下的 I/O 卸载方法。

- 30 在本申请实施例中，宿主机的 CPU 中增配了队列组件，且创新性地提出将队列组件用于 I/O 的处理过程中。基于此，可通过队列组件为宿主机上的实例提供虚拟 I/O 设备以及与虚拟 I/O 设备对应的设备队列；在 CPU 内，利用设备队列对实例与虚拟 I/O 设备之间发生的 I/O 请求进行调度；而原本的 I/O 卸载卡则只需负责监听设备队列中的 I/O 请求并传递 I/O 请求所相关的数据即可。这样，可将 I/O 处理过程中的设备队列上移到 CPU 内，从而在 I/O 调度过程中可与 CPU 内各部件联动，获取到 CPU 内各核心的实时负载，从而可
- 35

更加合理地实现 I/O 调度，而不再盲目调度，这可有效提高 I/O 的处理效率。

附图说明

此处所说明的附图用来提供对本申请的进一步理解，构成本申请的一部分，本申请的
5 示意性实施例及其说明用于解释本申请，并不构成对本申请的不当限定。在附图中：

图 1 为本申请一示例性实施例提供的云环境下的 I/O 卸载系统的结构示意图；

图 2 为本申请一示例性实施例提供的一种 I/O 卸载系统的可选实现方案对应的结构示
意图；

图 3 为本申请一示例性实施例提供的一种 I/O 卸载系统的可选实现方案的结构示意图；

10 图 4 为本申请一示例性实施例提供的一种实例创建过程的逻辑示意图；

图 5 为本申请一示例性实施例提供的一种实例销毁过程的逻辑示意图；

图 6 为本申请另一示例性实施例提供的一种云环境下的 I/O 卸载方法的流程示意图；

图 7 为本申请另一示例性实施例提供的另一种云环境下的 I/O 卸载方法的流程示意图；

图 8 为本申请又一示例性实施例提供的一种处理器 CPU 的结构示意图；

15 图 9 为本申请有一示例性实施例提供的一种 I/O 卸载卡的结构示意图。

具体实施方式

为使本申请的目的、技术方案和优点更加清楚，下面将结合本申请具体实施例及相应
的附图对本申请技术方案进行清楚、完整地描述。显然，所描述的实施例仅是本申请一部
20 分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做
出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

目前，I/O 的处理工作完全卸载到了 I/O 卸载卡中，对 I/O 的处理效率不佳。为此，本
申请的一些实施例中：宿主机的 CPU 中增配了队列组件，且创新性地提出将队列组件用于
I/O 卸载的处理过程中。基于此，可通过队列组件为宿主机上的实例提供虚拟 I/O 设备以及
25 与虚拟 I/O 设备对应的设备队列；在 CPU 内，利用设备队列对实例与虚拟 I/O 设备之间发
生的 I/O 请求进行调度；而原本的 I/O 卸载卡则只需负责监听设备队列中的 I/O 请求并传
递 I/O 请求所相关的数据即可。这样，可将 I/O 卸载处理过程中的设备队列上移到 CPU 内，
从而在 I/O 调度过程中可与 CPU 内各部件联动，获取到 CPU 内各核心的实时负载，从而
可更加合理地实现 I/O 调度，而不再盲目调度，这可有效提高 I/O 的处理效率。

30 以下结合附图，详细说明本申请各实施例提供的技术方案。

图 1 为本申请一示例性实施例提供的云环境下的 I/O 卸载系统的结构示意图。如图 1
所示，该系统包括：宿主机的 CPU 以及插接在宿主机上的 I/O 卸载卡。

其中，I/O 卸载卡可采用 DPU、IPU 等芯片进行实现，当然，本实施例并不限于。宿
主机可以是云环境中的物理机，参考图 1，在云环境中，单台宿主机上可运行多个实例。
35 云原生下的核心技术为虚拟化技术，基于虚拟化技术，可在宿主机的 CPU 之上虚拟出若干

VCPU，实例则可依托于 VCPU 而运行。

本实施例提出在宿主机的 CPU 中增配队列组件。其中，本实施例中可采用硬件方式或软件方式来实现队列组件。可选地，本实施例中，可采用专用集成电路 ASIC 等硬件来构建队列组件，当然，本实施例并不限于此。另外，本实施例中在 CPU 中增配的队列组件除了可在本实施例中用于为 I/O 处理工作提供队列功能外，也可复用到其它场景中，为其它场景提供队列功能。

本实施例中，队列组件可接入 CPU 的内部总线。本实施例并不限定队列组件接入 CPU 内部总线的形式，可根据实际情况来选用接入方式。例如，队列组件可与 CPU 内的 PCIe 控制器连接以接入 CPU 的内部总线。又例如，队列组件可通过 UCIe 总线连接到 CPU 其它组件。另外，队列组件还可与 CPU 内的内存控制器或者可选的加速模块等连接，以与 CPU 的内存或加速模块进行交互。而在另一侧，队列组件还可与插接在宿主机上的 I/O 卸载卡互联，可选地，队列组件可通过 CXL 等缓存一致性总线协议与 I/O 卸载卡进行互联。这样，参考图 1，CPU 中增配的队列组件可以支持 I/O 卸载卡与宿主机的 CPU 联动，改变传统 I/O 处理方案中 I/O 卸载卡与 CPU 数据链路分离的现状。

参考图 1，本实施例中的队列组件，可用于为宿主机上的实例提供虚拟 I/O 设备以及与虚拟 I/O 设备对应的设备队列。本实施例中，队列组件可对需要与实例进行 I/O 的物理 I/O 设备进行 I/O 虚拟化，以产生物理 I/O 设备对应的虚拟 I/O 设备（如图中的磁盘设备、网络设备）。虚拟出的磁盘设备可包括但不限于 blk 设备等，虚拟出的网络设备可包括但不限于 net 设备等。其中，队列组件可采用多种 I/O 虚拟化方案为宿主机上的实例提供给虚拟 I/O 设备，具体方案将在后文中详述。应当理解的是，这里的虚拟 I/O 设备是操作系统 OS 层面的定义，也即是，对宿主机的 OS 来说，其可见的是虚拟 I/O 设备，而该虚拟 I/O 设备代表的则是 I/O 过程中涉及到的各个实体设备。另外，本实施例中不限定设备队列的实现形式，可采用先入先出 FIFO 队列等实现形式。本实施例中，可利用设备队列来管理 I/O 请求。为便于描述，本实施例中，将从单个实例的角度来说明 I/O 卸载过程，但是应当理解的是，宿主机上可以承载多个实例，针对各个实例的 I/O 卸载方案一致。

值得说明的是，参考图 1，本实施例中，队列组件可为单个实例提供若干虚拟 I/O 设备，为便于描述，后文中将从单个虚拟 I/O 设备的角度来描述 I/O 卸载方案，但应当理解的是，可采用相同的优化方案来优化为单个实例提供的其它虚拟 I/O 设备所涉及到的 I/O 处理过程。

在此基础上，宿主机上的实例与队列组件为其提供的虚拟 I/O 设备之间可发生 I/O 请求。优选地，实例与虚拟 I/O 设备之间发生的 I/O 请求采用 io-uring 协议，通过采用如 io-uring 的通用协议，可避免 I/O 卸载卡与不同 CPU 平台联动时需要适配不同 CPU 平台实现的不同设备队列的问题，也即不再需要考虑适配性问题。当然，这仅是优选地，本实施例中实例与虚拟 I/O 设备之间发生的 I/O 请求还可采用协议，只需保证双方预先达成协议即可。

队列组件可在 CPU 内利用设备队列对实例与虚拟 I/O 设备之间发生的 I/O 请求进行调度。

其中，调度算法可以是五元组 hash、secret key hash 等，本实施例对此不做限定。这样，I/O 处理过程中涉及到的设备队列上移到了 CPU 内，而队列组件可与 CPU 内部的各组件联动，这使得队列组件基于设备队列所进行的 I/O 请求调度工作不再是盲目的，而是可以 CPU 内各核心的实时负载作为调度依据，从而可更加合理地完成 I/O 请求的调度，尤其是实例向虚拟 I/O 设备发出的读请求，这避免读请求过程中因调度不合理而导致的数据溢出问题。

正如前文提及的，队列组件还与插接在宿主机上的 I/O 卸载卡互联。这样，I/O 卸载卡将不再需要承担 I/O 请求调度的工作，这部分工作上移至 CPU 内，由队列组件承担。I/O 卸载卡的其它工作可保留，I/O 卸载卡可对接至队列组件为虚拟 I/O 设备提供的设备队列中，并监听设备队列中的 I/O 请求。据此，I/O 卸载卡可监听到实例与虚拟 I/O 设备之间发生的 I/O 请求。

这样，基于为 CPU 配置的队列组件，设备队列分别与虚拟 I/O 设备和 I/O 卸载卡对接，本实施例中的数据链路为：虚拟 I/O 设备将其与实例之间发生的 I/O 请求传递至设备队列，设备队列将 I/O 请求传递至 I/O 卸载卡，以使 I/O 卸载卡感知到 I/O 请求。

参考图 1，I/O 卸载卡还可用于在实例与虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。I/O 卸载卡可作为物理 I/O 设备与宿主机上的实例之间的中间件，为双方提供数据交互支持。

另外，本实施例中，队列组件中还可提供加速模块，加速模块可用于对实例与虚拟 I/O 设备之间发生的 I/O 请求所相关的数据执行加速处理。本实施例中的加速处理可包括但不限于加解密、压缩或统计卸载等。同样，I/O 卸载卡中也可提供加速模块，同样，其内的加速模块也可用于对实例与虚拟 I/O 设备之间发生的 I/O 请求所相关的数据执行加速处理。基于队列组件中提供的加速模块：

1、客户可配置指定虚拟 I/O 设备开启加解密、压缩、及统计卸载等功能，该虚拟 I/O 设备对应的设备队列绑定的 VF 可关联至指定加速模块。客户为该虚拟 I/O 设备分配的内存地址也可关联至该指定的加速模块，这样，读写该内存地址，数据都会经过指定的加速模块进行加速处理。

2、实例中的应用向虚拟 I/O 设备发起写请求时，数据进入客户分配的内存空间前，加速模块可对数据进行预处理。

3、I/O 卸载卡的 I/O 计算引擎从通信组件收到数据，I/O 计算引擎将数据 DMA 到客户预分配的内存空间。数据进入客户分配的内存空间前，加速模块可对数据进行预处理。

这样，基于队列组件中提供的加速模块，本实施例中的数据链路变为：虚拟 I/O 设备将其与实例之间发生的 I/O 请求传递至设备队列；设备队列触发加速模块访问内存地址并对 I/O 请求所相关的数据进行加速处理；设备队列将 I/O 请求传递至 I/O 卸载卡，以使 I/O 卸载卡感知到 I/O 请求并从内存地址读取加速处理后的数据。这可使加速模块的加速范围覆盖至数据的全部路径，比如，队列组件提供的加解密模块可保证数据在全路径的加密状态，改变传统方案中只能由 I/O 卸载卡在数据传输路径的末端进行数据加速的现状。

综上，本实施例中，宿主机的 CPU 中增配了队列组件，且创新性地提出将队列组件用于 I/O 的处理过程中。基于此，可通过队列组件为宿主机上的实例提供虚拟 I/O 设备以及
5 与虚拟 I/O 设备对应的设备队列；在 CPU 内，利用设备队列对实例与虚拟 I/O 设备之间发生的 I/O 请求进行调度；而原本的 I/O 卸载卡则只需负责监听设备队列中的 I/O 请求并传递 I/O 请求所相关的数据即可。这样，可将 I/O 处理过程中的设备队列上移到 CPU 内，从而在 I/O 调度过程中可与 CPU 内各部件联动，获取到 CPU 内各核心的实时负载，从而可更加合理地实现 I/O 调度，而不再盲目调度，这可有效提高 I/O 的处理效率。

图 2 为本申请一示例性实施例提供的一种 I/O 卸载系统的可选实现方案对应的结构示意图。参考图 2，在上述或下述实施例中，队列组件提供的设备队列可包括第一层队列和
10 第二层队列。其中，第一层队列可对接实例下的 VCPU，第二层队列可通过 I/O 卸载卡而对接至物理 I/O 设备。参考图 2，对于为实例提供的其中一个虚拟 I/O 设备来说，队列组件可为虚拟 I/O 设备关联第一指定数量的第一层队列和第二指定数量的第二层队列；建立虚拟 I/O 设备关联的第一层队列和第二层队列之间的映射关系；将虚拟 I/O 设备关联的第一层队列对接至实例下的各个 VCPU；将虚拟 I/O 设备关联的第二层队列通过 I/O 卸载卡对接至虚拟 I/O 设备对应的物理 I/O 设备。
15

图 3 为本申请一示例性实施例提供的一种 I/O 卸载系统的可选实现方案的结构示意图，其中示出了将虚拟 I/O 设备关联的第一层队列对接至实例下的 VCPU 的实现方案以及将虚拟 I/O 设备关联的第二层队列对接至虚拟 I/O 设备对应的物理 I/O 设备的实现方案。

参考图 3，正如前文提及的，队列组件可采用多种 I/O 虚拟化方案为宿主机上的实例
20 提供给虚拟 I/O 设备。在一种示例性的实现方案中：可采用 SRIOV (Single Root I/O Virtualization) 技术，为需要与实例进行 I/O 的物理 I/O 设备创建 VF 设备，VF 设备用于与实例下的各个 VCPU 进行数据交换；利用实例的操作系统将 VF 设备注册为指定类型的虚拟 I/O 设备。例如，实例的操作系统可根据 VF 的 PCIe <vender ID、device ID>，将 VF 注册为磁盘、网络设备等。其中，本实施例中的物理 I/O 设备可包括但不限于云网卡、云盘
25 (例如固态硬盘 SSD) 等高性能的 I/O 设备。实际应用中，每个 SR-I/OV 设备都可有一个物理功能 (Physical FunctI/On, PF)，并且每个 PF 最多可有 64000 个与其关联的虚拟功能 (Virtual FunctI/On, VF)。PF 可以通过寄存器创建 VF，这些寄存器设计有专用于此目的的属性。一旦在 PF 中启用了 SR-I/OV，就可以通过 PF 的总线、设备和功能编号等访问各个 VF 的 PCI 配置空间。每个 VF 都具有一个 PCI 内存空间，用于映射其寄存器集。VF 设备驱动程序对寄存器集进行操作以启用其功能，并且显示为实际存在的 PCI
30 设备。创建 VF 后，可以直接将其指定给实例中的各个应用程序，这样，队列组件中的 VF 设备可与实例下各个 VCPU 进行数据交换。

基于此，参考图 3，可将虚拟 I/O 设备关联的第一层队列绑定至虚拟 I/O 设备对应的 VF 设备，以与实例下各个 VCPU 进行数据交换。

35 继续参考图 3，对于 I/O 卸载卡来说，可对物理 I/O 设备进行模拟，以产生物理 I/O 设

备对应的模拟设备；基于此，可将虚拟 I/O 设备关联的第二层队列绑定至模拟设备，以将虚拟 I/O 设备关联的第二层队列对接至虚拟 I/O 设备对应的物理 I/O 设备。其中，I/O 卸载卡可采用软件模拟或硬件模拟的方式来实现对物理 I/O 设备的模拟，本实施例在此不做限定。应当理解的是，在云环境中，物理 I/O 设备的部署形式可能是分布式或集群式的，I/O 卸载卡是在相应的分布式系统或集群之上模拟出模拟设备以参与到 I/O 处理过程中，最终，I/O 卸载卡可通过模拟设备来将 I/O 请求准确传递至物理 I/O 设备。也即是，I/O 卸载卡中的 I/O 计算引擎可提供指定数量个模拟设备，每个模拟设备与 CPU 内的设备队列中的指定数量个第二层队列绑定以读取或发送数据。当然，本实施例中，还可采用其它实现方案来将虚拟 I/O 设备关联的第二层队列对接至虚拟 I/O 设备对应的物理 I/O 设备，而并不限于模拟这一种方案，在此不再穷举。图 4 为本申请一示例性实施例提供的一种实例创建过程的逻辑示意图。图 5 为本申请一示例性实施例提供的一种实例销毁过程的逻辑示意图。以下将通过实例的创建/销毁过程来说明上述的队列对接方案。

参考图 4，一种示例性的实例创建过程可以是：

- 1、客户通过控制台、或 openAPI 发起实例创建请求。
- 2、控制台可将创建请求调度到最合适的宿主机。
- 3、宿主机上可运行实例管控 agent，以在 I/O 卸载卡的 I/O 计算引擎上创建云网卡或云盘等物理 I/O 设备对应的模拟设备，并绑定至设备队列中指定数量个第二层队列。
- 4、宿主机的 CPU 中的队列组件可在设备队列上创建所需的多个 VF，并绑定指定的第一层队列，并建立与步骤 3 中的第二层队列之间的映射关系。
- 5、宿主机上的实例管控 agent 可通过调用 Hypervisor，创建实例。

参考图 5，一种示例性的实例销毁过程可以是：

- 1、客户通过控制台、或 openAPI 发起实例销毁请求。
- 2、控制台可查找到该实例所在宿主机，通过该宿主机上实例管控 agent 下发删除命令。
- 3、实例管控 agent，可在 I/O 卸载卡的 I/O 计算引擎上删除相关的云网卡或云盘等物理 I/O 设备对应的模拟设备，宿主机的 CPU 中的队列组件可在设备队列上删除相关的 VF。
- 4、实例管控 agent 调用 Hypervisor，删除实例。

从图 4 和图 5 可知，本实施例中，在实例的创建/销毁过程中，将从单实例的粒度来创建所需的多个虚拟 I/O 设备及多个模拟设备并进行两层队列的对接过程。值得说明的是，在实例正常运行的过程中，如果发生物理 I/O 设备的新增/减少，则可从单设备的粒度来配置图 3 中单个虚拟 I/O 设备所管理的一串实体对象（VF 设备-第一层队列-第二层队列-模拟设备），也即是，每增加一个需要与实例进行 I/O 的物理 I/O 设备，则可增加图 3 中的 1 个虚拟 I/O 设备及该虚拟 I/O 设备所管理的一串实体对象。

另外，值得说明的是，本实施例中，虚拟 I/O 设备是通过对物理 I/O 设备进行虚拟化而产生的，其主要用于支持宿主机上的实例发现物理 I/O 设备。而 I/O 卸载卡中的模拟设备则是通过对物理 I/O 设备进行模拟而产生的，其主要用于模拟物理 I/O 设备的硬件行为。

而基于全文中提供的各个可选实现方案可知，本实施例中，参与到 I/O 处理过程中的实体对象可能包括：VF 设备-第一层队列-第二层队列-模拟设备-网卡-物理 I/O 设备。而前述的虚拟 I/O 设备所代表的则是参与到 I/O 处理过程中的这些实体。而 I/O 卸载卡中的模拟设备则代表的是这些实体中最后提到的物理 I/O 设备。本实施例中的虚拟 I/O 设备和模拟设备功能均用于支持云原生下的 I/O 虚拟化。这样，本实施例中，虚拟 I/O 设备关联的第一层队列可绑定至 VF 设备与实例下的各个 VCPU 进行数据交换；虚拟 I/O 设备关联的第二层队列可绑定至 I/O 卸载卡中的模拟设备并进行数据交换。虚拟 I/O 设备关联的第一层队列可 N:1 映射至其关联的第二层队列，也即是，虚拟 I/O 设备关联的第一层队列的数量通常大于其关联的第二层队列的数量。基于这种双层的队列结构，可通过从第一层队列向第二层队列映射而将实例向虚拟 I/O 设备发起的写请求调度至合适的物理 I/O 设备；可通过从第二层队列向第一层队列映射而将实例向虚拟 I/O 设备发起的读请求调度至合适的 VCPU。

以下将对基于双层的队列结构进行调度的过程进行详述。

参考图 2，队列组件可在实例向虚拟 I/O 设备发起读请求的过程中，读取实例下的各个 VCPU 的负载信息；根据负载信息将读请求调度至虚拟 I/O 设备关联的第一层队列中的第一队列，以使用第一队列所对接的 VCPU 处理读请求。可选地，第一队列可以是负载最低的 VCPU 对接的队列。这样，可将读请求调度至负载最优的 VCPU 上进行处理，从而可提高 I/O 处理效率。在此过程中，I/O 卸载卡可获取读请求对应的响应报文；将响应报文中的元数据信息加入虚拟 I/O 设备关联的第二层队列中的指定队列。而队列组件则可根据负载信息将指定队列中的元数据信息调度至第一队列中。另外，宿主机可为实例分配内存空间，I/O 卸载卡可基于内存空间来在实例与虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。在一种示例性方案中：I/O 卸载卡可将响应报文中的数据部分写入实例对应的内存空间；将数据部分所在的内存地址添加至元数据信息中；基于此，队列组件可触发目标队列对接的 VCPU 按照元数据信息中的内存地址读取响应报文中的数据部分。从而完成实例向虚拟 I/O 设备发起的读请求。

其中，本实施例中，实例读写虚拟 I/O 设备时，系统调用可分配类似网络的 sk_buf 等内存空间。设备队列通过地址访问引用该内存空间，I/O 卸载卡可通过 DMA、或者类似 Intel SVM 模式访问该内存空间。

举例来说，一种实际的应用方案可以是：

1、实例中的应用可基于 io-uring 协议打开虚拟 I/O 设备（如图 2 中磁盘 设备、网络设备），使用 io_uring_smp_store_release 提交读请求，使用 io_uring_enter 触发接收方到设备队列中取数据。

2、I/O 卸载卡的 I/O 计算引擎从通信组件（如图 2 中的网卡）收到响应报文，基于报文元数据部分，确定报文所属的模拟设备。将报文的数据部分写入该模拟设备的 DMA 内存空间。数据写入过程中，可以使用 I/O 卸载卡的加速模块进行数据预处理。

3、I/O 计算引擎生成 Io-uring 格式的读请求的元数据信息，元数据信息中包含数据部分写入的 DMA 内存地址。I/O 计算引擎可将读请求的元数据信息写入到 CPU 内设备队列的第二层队列中。

4、队列组件中运行的队列调度器，可读取实例各 VCPU 的功耗、时间片利用率、及 PMU 等，实时统计个 VCPU 的实时负载，结合设定的五元组 hash、secret key hash 等调度策略，将第二层队列中读请求调度到第一层队列中的目标队列。

5、实例的操作系统可唤醒提交读请求的应用，从设备队列的目标队列中读取读请求的元数据信息，从元数据信息中读取数据部分所在的内存地址，从该内存地址读取数据。

参考图 2，队列组件可将实例向虚拟 I/O 设备发起的写请求的元数据信息加入虚拟 I/O 设备关联的第一层队列中；将写请求的元数据信息调度至虚拟 I/O 设备关联的第二层队列中的第二队列；而 I/O 卸载卡，则可利用第二队列对接的物理 I/O 设备处理写请求。在此过程中，队列组件可将写请求对应的数据部分写入实例对应的内存空间；将数据部分所在的内存地址添加至元数据信息中；I/O 卸载卡可从第二队列中读取元数据信息；按照元数据信息中的内存地址获取写请求的数据部分；将数据部分发送至第二队列对接的物理 I/O 设备中。

举例来说，一种实际的应用方案可以是：

1、实例中的应用可基于 io-uring 协议打开虚拟 I/O 设备（如图 2 中磁盘设备、网络设备），使用 `io_uring_smp_store_release` 提交写请求，使用 `io_uring_enter` 触发接收方到设备队列中取数据。

2、写请求的数据部分在为应用分配的内存空间中，写请求的元数据信息进入设备队列的第一层队列中。

3、队列组件中运行的队列调度器，可根据第一层队列与第二层队列之间的映射关系，第一层队列的权重，及第二层队列的空闲队列条目数量等信息，将第一层队列中的写请求送到指定的第二层队列中。

4、使用 CPU 内的加速模块进行数据预处理，比如数据加密、数据压缩等。

5、I/O 卸载卡的 I/O 计算引擎从 CPU 内的队列组件的第二层队列中读取写请求的元数据信息，以获得数据部分的访存地址等信息。

6、I/O 计算引擎增加虚拟网卡、云盘等元数据报文头，将应用提交的数据，通过 I/O 卸载卡的通信组件发送到目标宿主机、或者后端存储集群。数据发送过程中，可以使用 I/O 卸载卡的加速模块进行数据预处理。

据此，可通过队列组件实现对实例与虚拟 I/O 设备之间发生的 I/O 请求的合理调度。设备队列上移到 CPU 内，可支持在实例上使用工具带内配置 CPU 内设备队列的调度策略，而不再需要像传统方案那样需要由控制台进行配置，本实施例提供的配置设备队列中调度策略的方案可使得配置 QPS 提升若干个数量级，解决了传统方案实时性不高，不能满足容器、serverless 场景的快速创建销毁场景需求的问题。且，I/O 卸载卡传递 I/O 请求所相关

的数据的过程中,可使用 CPU 内的加解密模块对数据进行加密后再进入 I/O 卸载卡,这样, I/O 卸载卡将无法再看到数据明文,这使得用户数据的安全性大幅提升。

图 6 为本申请另一示例性实施例提供的一种云环境下的 I/O 卸载方法的流程示意图,该方法可由前述系统实施例中的宿主机中的 CPU 实施,该 CPU 中装配有队列组件。参考

5 图 6,该方法可包括:

步骤 600、利用队列组件为宿主机上的实例提供虚拟 I/O 设备;

步骤 601、利用队列组件为虚拟 I/O 设备配置对应的设备队列;

步骤 602、在 CPU 内,利用设备队列对实例与虚拟 I/O 设备之间发生的 I/O 请求进行调度,以供宿主机上插接的 I/O 卸载卡监听设备队列中的 I/O 请求并在实例与虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

在一可选实施例中,采用专用集成电路 ASIC 构建队列组件,队列组件接入 CPU 的内部总线。

在一可选实施例中,设备队列分别与虚拟 I/O 设备和 I/O 卸载卡对接,虚拟 I/O 设备将其与实例之间发生的 I/O 请求传递至设备队列,设备队列将 I/O 请求传递至 I/O 卸载卡,以使 I/O 卸载卡感知到 I/O 请求。

在一可选实施例中,设备队列包括第一层队列和第二层队列,所述方法还包括:
为虚拟 I/O 设备关联第一指定数量的第一层队列和第二指定数量的第二层队列;
建立虚拟 I/O 设备关联的第一层队列和第二层队列之间的映射关系;
将虚拟 I/O 设备关联的第一层队列对接至实例下的各个 VCPU;
20 将虚拟 I/O 设备关联的第二层队列通过 I/O 卸载卡对接至虚拟 I/O 设备对应的物理 I/O 设备。

在一可选实施例中,步骤对实例与虚拟 I/O 设备之间发生的 I/O 请求进行调度,包括:
在实例向虚拟 I/O 设备发起读请求的过程中,读取实例下的各个 VCPU 的负载信息;
根据负载信息将读请求调度至虚拟 I/O 设备关联的第一层队列中的第一队列,以使用
25 第一队列所对接的 VCPU 处理读请求。

在一可选实施例中,步骤根据负载信息将读请求调度至虚拟 I/O 设备关联的第一层队列中的第一队列,包括:

根据负载信息将指定队列中的元数据信息调度至第一队列中;其中,由 I/O 卸载卡获取读请求对应的响应报文并将响应报文中的元数据信息加入虚拟 I/O 设备关联的第二层队列中的指定队列。

30 在一可选实施例中,该方法还包括:

触发目标队列对接的 VCPU 按照元数据信息中的内存地址读取响应报文中的数据部分;其中,由 I/O 卸载卡将响应报文中的数据部分写入实例对应的内存空间;将数据部分所在的内存地址添加至元数据信息中。

35 在一可选实施例中,步骤利用设备队列对实例与虚拟 I/O 设备之间发生的 I/O 请求进

行调度, 包括:

将实例向虚拟 I/O 设备发起的写请求的元数据信息加入虚拟 I/O 设备关联的第一层队列中;

将写请求的元数据信息调度至虚拟 I/O 设备关联的第二层队列中的第二队列;

5 其中, 由 I/O 卸载卡利用第二队列对接的物理 I/O 设备处理写请求。

在一可选实施例中, 该方法还包括:

将写请求对应的数据部分写入实例对应的内存空间; 将数据部分所在的内存地址添加至元数据信息中;

10 其中, 由 I/O 卸载卡从第二队列中读取元数据信息; 按照元数据信息中的内存地址获取写请求的数据部分, 并将数据部分发送至第二队列对接的物理 I/O 设备中。

在一可选实施例中, 步骤为宿主机上的实例提供虚拟 I/O 设备, 包括:

对需要与实例进行 I/O 的物理 I/O 设备进行 I/O 虚拟化, 以产生物理 I/O 设备对应的虚拟 I/O 设备。

15 在一可选实施例中, 步骤对需要与实例进行 I/O 的物理 I/O 设备进行 I/O 虚拟化, 包括:

采用 SRIOV 技术, 为需要与实例进行 I/O 的物理 I/O 设备创建 VF 设备, VF 设备用于与实例下的各个 VCPU 进行数据交换;

利用实例的操作系统将 VF 设备注册为指定类型的虚拟 I/O 设备。

在一可选实施例中, 实例与虚拟 I/O 设备之间发生的 I/O 请求采用 io-uring 协议。

20 在一可选实施例中, 该方法还包括:

利用加速模块对实例与虚拟 I/O 设备之间发生的 I/O 请求所相关的数据执行加速处理, 加速处理包括: 加解密、压缩和统计卸载中的一种或多种处理。

25 在一可选实施例中, 加速模块绑定至宿主机为虚拟 I/O 设备分配的内存地址, 虚拟 I/O 设备将其与实例之间发生的 I/O 请求传递至设备队列; 设备队列触发加速模块访问内存地址并对 I/O 请求所相关的数据进行加速处理; 设备队列将 I/O 请求传递至 I/O 卸载卡, 以使 I/O 卸载卡感知到 I/O 请求并从内存地址读取加速处理后的数据。

值得说明的是, 上述关于 I/O 卸载方法各实施例中的技术细节, 可参考前述的系统实施例中关于 CPU 的相关描述, 为节省篇幅, 在此不再赘述, 但这不应造成本申请保护范围的损失。

30 图7为本申请另一示例性实施例提供的另一种云环境下的 I/O 卸载方法的流程示意图, 该方法可由前述系统实施例中插接在宿主机上的 I/O 卸载卡实施, 宿主机的 CPU 中装配有队列组件。参考图 7, 该方法可包括:

步骤 700、从队列组件中的设备队列中监听宿主机上的实例与队列组件为实例提供的虚拟 I/O 设备之间发生的 I/O 请求;

35 步骤 701、获取监听到 I/O 请求所相关的数据;

步骤 702、在实例与虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

在一可选实施例中，采用专用集成电路 ASIC 构建队列组件，队列组件接入 CPU 的内部总线。

5 在一可选实施例中，设备队列包括第一层队列和第二层队列，且由队列组件为虚拟 I/O 设备关联第一指定数量的第一层队列和第二指定数量的第二层队列；建立虚拟 I/O 设备关联的第一层队列和第二层队列之间的映射关系；将虚拟 I/O 设备关联的第一层队列对接至实例下的各个 VCPU；将虚拟 I/O 设备关联的第二层队列通过 I/O 卸载卡对接至虚拟 I/O 设备对应的物理 I/O 设备。

10 在一可选实施例中，该方法还可包括：

对物理 I/O 设备进行模拟，以产生物理 I/O 设备对应的模拟设备；

将虚拟 I/O 设备关联的第二层队列绑定至模拟设备，以将虚拟 I/O 设备关联的第二层队列对接至物理 I/O 设备。

在一可选实施例中，该方法还可包括：

15 获取读请求对应的响应报文；将响应报文中的元数据信息加入虚拟 I/O 设备关联的第二层队列中的指定队列，以供队列组件根据负载信息将指定队列中的元数据信息调度至第一队列中。

在一可选实施例中，步骤在实例与虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据，可包括：将响应报文中的数据部分写入实例对应的内存空间；

20 将数据部分所在的内存地址添加至元数据信息中，以供队列组件触发目标队列对接的 VCPU 按照元数据信息中的内存地址读取响应报文中的数据部分。

在一可选实施例中，可由队列组件将实例向虚拟 I/O 设备发起的写请求的元数据信息加入虚拟 I/O 设备关联的第一层队列中；将写请求的元数据信息调度至虚拟 I/O 设备关联的第二层队列中的第二队列；该方法还包括：

25 利用第二队列对接的物理 I/O 设备处理写请求。

在一可选实施例中，步骤在实例与虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据，可包括：从第二队列中读取元数据信息；按照元数据信息中的内存地址获取写请求的数据部分；将数据部分发送至第二队列对接的物理 I/O 设备中；

30 其中，由队列组件将写请求对应的数据部分写入实例对应的内存空间；将数据部分所在的内存地址添加至元数据信息中。

在一可选实施例中，实例与虚拟 I/O 设备之间发生的 I/O 请求采用 io-uring 协议。

值得说明的是，上述关于 I/O 卸载方法各实施例中的技术细节，可参考前述的系统实施例中关于 I/O 卸载卡的相关描述，为节省篇幅，在此不再赘述，但这不应造成本申请保护范围的损失。

35 另外，在上述实施例及附图中的描述的一些流程中，包含了按照特定顺序出现的多个

操作,但是应该清楚了解,这些操作可以不按照其在本文中出现的顺序来执行或并行执行,操作的序号如 700、701 等,仅仅是用于区分开各个不同的操作,序号本身不代表任何的执行顺序。另外,这些流程可以包括更多或更少的操作,并且这些操作可以按顺序执行或并行执行。需要说明的是,本文中的“第一”、“第二”等描述,是用于区分不同的队列层、

5 队列等,不代表先后顺序,也不限定“第一”和“第二”是不同的类型。

图 8 为本申请又一示例性实施例提供的一种处理器 CPU 的结构示意图,该 CPU 安装在宿主机中,该 CPU 内装配有队列组件 80。如图 8 所示,该 CPU 可用于执行所述一条或多条计算机指令,以用于:

利用所述队列组件 80 为所述宿主机上的实例提供虚拟 I/O 设备;

10 利用所述队列组件 80 为所述虚拟 I/O 设备配置对应的设备队列;

在所述 CPU 内,利用所述设备队列对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求进行调度,以供所述宿主机上插接的 I/O 卸载卡监听所述设备队列中的 I/O 请求并在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

在一可选实施例中,采用专用集成电路 ASIC 构建队列组件 80,队列组件 80 接入 CPU

15 的内部总线。

在一可选实施例中,设备队列分别与虚拟 I/O 设备和 I/O 卸载卡对接,虚拟 I/O 设备将其与实例之间发生的 I/O 请求传递至设备队列,设备队列将 I/O 请求传递至 I/O 卸载卡,以使 I/O 卸载卡感知到 I/O 请求。

在一可选实施例中,设备队列包括第一层队列和第二层队列,该队列组件 80 还可用于:

20

为虚拟 I/O 设备关联第一指定数量的第一层队列和第二指定数量的第二层队列;

建立虚拟 I/O 设备关联的第一层队列和第二层队列之间的映射关系;

将虚拟 I/O 设备关联的第一层队列对接至实例下的各个 VCPU;

将虚拟 I/O 设备关联的第二层队列通过 I/O 卸载卡对接至虚拟 I/O 设备对应的物理 I/O

25 设备。

在一可选实施例中,队列组件 80 在对实例与虚拟 I/O 设备之间发生的 I/O 请求进行调度的过程中,可用于:

在实例向虚拟 I/O 设备发起读请求的过程中,读取实例下的各个 VCPU 的负载信息;

根据负载信息将读请求调度至虚拟 I/O 设备关联的第一层队列中的第一队列,以使用

30 第一队列所对接的 VCPU 处理读请求。

在一可选实施例中,队列组件 80 在根据负载信息将读请求调度至虚拟 I/O 设备关联的第一层队列中的第一队列的过程中,可用于:

根据负载信息将指定队列中的元数据信息调度至第一队列中;其中,由 I/O 卸载卡获取读请求对应的响应报文并将响应报文中的元数据信息加入虚拟 I/O 设备关联的第二层队列中的指定队列。

35

在一可选实施例中，队列组件 80 还可用于：

触发目标队列对接的 VCPU 按照元数据信息中的内存地址读取响应报文中的数据部分；其中，由 I/O 卸载卡将响应报文中的数据部分写入实例对应的内存空间；将数据部分所在的内存地址添加至元数据信息中。

- 5 在一可选实施例中，队列组件 80 在利用设备队列对实例与虚拟 I/O 设备之间发生的 I/O 请求进行调度的过程中，可用于：

将实例向虚拟 I/O 设备发起的写请求的元数据信息加入虚拟 I/O 设备关联的第一层队列中；

将写请求的元数据信息调度至虚拟 I/O 设备关联的第二层队列中的第二队列；

- 10 其中，由 I/O 卸载卡利用第二队列对接的物理 I/O 设备处理写请求。

在一可选实施例中，队列组件 80 还可用于：

将写请求对应的数据部分写入实例对应的内存空间；将数据部分所在的内存地址添加至元数据信息中；

- 15 其中，由 I/O 卸载卡从第二队列中读取元数据信息；按照元数据信息中的内存地址获取写请求的数据部分，并将数据部分发送至第二队列对接的物理 I/O 设备中。

在一可选实施例中，队列组件 80 在为宿主机上的实例提供虚拟 I/O 设备的过程中，可用于：

对需要与实例进行 I/O 的物理 I/O 设备进行 I/O 虚拟化，以产生物理 I/O 设备对应的虚拟 I/O 设备。

- 20 在一可选实施例中，队列组件 80 在对需要与实例进行 I/O 的物理 I/O 设备进行 I/O 虚拟化的过程中，可用于：

采用 SRIOV 技术，为需要与实例进行 I/O 的物理 I/O 设备创建 VF 设备，VF 设备用于与实例下的各个 VCPU 进行数据交换；

利用实例的操作系统将 VF 设备注册为指定类型的虚拟 I/O 设备。

- 25 在一可选实施例中，实例与虚拟 I/O 设备之间发生的 I/O 请求采用 io-uring 协议。

在一可选实施例中，队列组件 80 还可包括加速模块，队列组件 80 还可用于：

利用加速模块对实例与虚拟 I/O 设备之间发生的 I/O 请求所相关的数据执行加速处理，加速处理包括：加解密、压缩和统计卸载中的一种或多种处理。

- 30 在一可选实施例中，加速模块绑定至宿主机为虚拟 I/O 设备分配的内存地址，虚拟 I/O 设备将其与实例之间发生的 I/O 请求传递至设备队列；设备队列触发加速模块访问内存地址并对 I/O 请求所相关的数据进行加速处理；设备队列将 I/O 请求传递至 I/O 卸载卡，以使 I/O 卸载卡感知到 I/O 请求并从内存地址读取加速处理后的数据。

- 35 值得说明的是，上述关于 CPU 各实施例中的技术细节，可参考前述的系统实施例中关于 CPU 的相关描述，为节省篇幅，在此不再赘述，但这不应造成本申请保护范围的损失。

图 9 为本申请有一示例性实施例提供的一种 I/O 卸载卡的结构示意图, 该 I/O 卸载卡插接于宿主机上, 宿主机的 CPU 内装配有队列组件。参考图 9, 该 I/O 卸载卡可包括存储器 90 和处理器 91, 存储器 90 用于存储一条或多条计算机指令; 处理器 91 与存储器 90 耦合, 用于执行一条或多条计算机指令, 以用于:

- 5 从队列组件中的设备队列中监听宿主机上的实例与队列组件为实例提供的虚拟 I/O 设备之间发生的 I/O 请求;

获取监听到的 I/O 请求所相关的数据;

在实例与虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

- 10 在一可选实施例中, 采用专用集成电路 ASIC 构建队列组件, 队列组件接入 CPU 的内部总线。

在一可选实施例中, 设备队列包括第一层队列和第二层队列, 且由队列组件为虚拟 I/O 设备关联第一指定数量的第一层队列和第二指定数量的第二层队列; 建立虚拟 I/O 设备关联的第一层队列和第二层队列之间的映射关系; 将虚拟 I/O 设备关联的第一层队列对接至
15 实例下的各个 VCPU; 将虚拟 I/O 设备关联的第二层队列通过 I/O 卸载卡对接至虚拟 I/O 设备对应的物理 I/O 设备。

在一可选实施例中, 处理器 91 还可用于:

对物理 I/O 设备进行模拟, 以产生物理 I/O 设备对应的模拟设备;

- 20 将虚拟 I/O 设备关联的第二层队列绑定至模拟设备, 以将虚拟 I/O 设备关联的第二层队列对接至物理 I/O 设备。

在一可选实施例中, 处理器 91 还可用于:

获取读请求对应的响应报文; 将响应报文中的元数据信息加入虚拟 I/O 设备关联的第二层队列中的指定队列, 以供队列组件根据负载信息将指定队列中的元数据信息调度至第一队列中。

- 25 在一可选实施例中, 处理器 91 在实例与虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据的过程中, 可用于: 将响应报文中的数据部分写入实例对应的内存空间; 将数据部分所在的内存地址添加至元数据信息中, 以供队列组件触发目标队列对接的 VCPU 按照元数据信息中的内存地址读取响应报文中的数据部分。

- 30 在一可选实施例中, 可由队列组件将实例向虚拟 I/O 设备发起的写请求的元数据信息加入虚拟 I/O 设备关联的第一层队列中; 将写请求的元数据信息调度至虚拟 I/O 设备关联的第二层队列中的第二队列; 处理器 91 还可用于:

利用第二队列对接的物理 I/O 设备处理写请求。

- 35 在一可选实施例中, 处理器 91 在实例与虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据的过程中, 可用于: 从第二队列中读取元数据信息; 按照元数据信息中的内存地址获取写请求的数据部分; 将数据部分发送至第二队列对接的物理

I/O 设备中；

其中，由队列组件将写请求对应的数据部分写入实例对应的内存空间；将数据部分所在的内存地址添加至元数据信息中。

在一可选实施例中，实例与虚拟 I/O 设备之间发生的 I/O 请求采用 io-uring 协议。

5 进一步，如图 9 所示，该 I/O 卸载卡还包括：通信组件 92、电源组件 93 等其它组件。图 9 中仅示意性给出部分组件，并不意味着 I/O 卸载卡只包括图 9 所示组件。

值得说明的是，上述关于 I/O 卸载卡各实施例中的技术细节，可参考前述的系统实施例中关于 I/O 卸载卡的相关描述，为节省篇幅，在此不再赘述，但这不应造成本申请保护范围的损失。

10 相应地，本申请实施例还提供一种存储有计算机程序的计算机可读存储介质，计算机程序被执行时能够实现上述方法实施例中可由 CPU 或 I/O 卸载卡执行的各步骤。

上述图 9 中的存储器，用于存储计算机程序，并可被配置为存储其它各种数据以支持在计算平台上的操作。这些数据的示例包括用于在计算平台上操作的任何应用程序或方法的指令，联系人数据，电话簿数据，消息，图片，视频等。存储器可以由任何类型的易失性或非易失性存储设备或者它们的组合实现，如静态随机存取存储器（SRAM），电可擦除可编程只读存储器（EEPROM），可擦除可编程只读存储器（EPROM），可编程只读存储器（PROM），只读存储器（ROM），磁存储器，快闪存储器，磁盘或光盘。

上述图 9 中的通信组件，被配置为便于通信组件所在设备和其他设备之间有线或无线方式的通信。通信组件所在设备可以接入基于通信标准的无线网络，WIFI、2G、3G、4G/LTE、5G 等移动通信网络，或它们的组合。在一个示例性实施例中，通信组件经由广播信道接收来自外部广播管理系统的广播信号或广播相关信息。在一个示例性实施例中，所述通信组件还包括近场通信（NFC）模块，以促进短程通信。例如，在 NFC 模块可基于射频识别（RFID）技术，红外数据协会（IrDA）技术，超宽带（UWB）技术，蓝牙（BT）技术和其他技术来实现。

25 上述图 9 中的电源组件，为电源组件所在设备的各种组件提供电力。电源组件可以包括电源管理系统，一个或多个电源，及其他与为电源组件所在设备生成、管理和分配电力相关联的组件。

本领域内的技术人员应明白，本申请的实施例可提供为方法、系统、或计算机程序产品。因此，本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

30 本申请是参照根据本申请实施例的方法、设备（系统）、和计算机程序产品的流程图和 / 或方框图来描述的。应理解可由计算机程序指令实现流程图和 / 或方框图中的每一流

程和 / 或方框、以及流程图和 / 或方框图中的流程和 / 或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

在一个典型的配置中，计算设备包括一个或多个处理器 (CPU)、输入/输出接口、网络接口和内存。

内存可能包括计算机可读介质中的非永久性存储器，随机存取存储器 (RAM) 和/或非易失性内存等形式，如只读存储器 (ROM) 或闪存(flash RAM)。内存是计算机可读介质的示例。

计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括，但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、其他类型的随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器 (CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带，磁带式磁盘存储或其他磁性存储设备或任何其他非传输介质，可用于存储可以被计算设备访问的信息。按照本文中的界定，计算机可读介质不包括暂存电脑可读媒体 (transitory media)，如调制的数据信号和载波。

还需要说明的是，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个.....”限定的要素，并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

以上所述仅为本申请的实施例而已，并不用于限制本申请。对于本领域技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

权 利 要 求 书

1、一种云环境下的 I/O 卸载系统，包括：宿主机的 CPU 以及插接在所述宿主机上的 I/O 卸载卡，所述 CPU 内装配有队列组件；

5 所述队列组件，用于为所述宿主机上的实例提供虚拟 I/O 设备以及与所述虚拟 I/O 设备对应的设备队列；在所述 CPU 内，利用所述设备队列对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求进行调度；

所述 I/O 卸载卡，用于监听所述设备队列中的 I/O 请求；在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

10 2、根据权利要求 1 所述的系统，采用专用集成电路 ASIC 构建所述队列组件，所述队列组件接入所述 CPU 的内部总线。

3、根据权利要求 1 所述的系统，所述设备队列分别与所述虚拟 I/O 设备和所述 I/O 卸载卡对接，所述虚拟 I/O 设备将其与所述实例之间发生的 I/O 请求传递至所述设备队列，所述设备队列将所述 I/O 请求传递至所述 I/O 卸载卡，以使所述 I/O 卸载卡感知到所述 I/O 请求。

15 4、根据权利要求 1 所述的系统，所述设备队列包括第一层队列和第二层队列，所述队列组件还用于：

为所述虚拟 I/O 设备关联第一指定数量的第一层队列和第二指定数量的第二层队列；

20 建立所述虚拟 I/O 设备关联的第一层队列和第二层队列之间的映射关系；

将所述虚拟 I/O 设备关联的第一层队列对接至所述实例下的各个 VCPU；

将所述虚拟 I/O 设备关联的第二层队列通过 I/O 卸载卡对接至所述虚拟 I/O 设备对应的物理 I/O 设备。

5 5、根据权利要求 4 所述的系统，所述 I/O 卸载卡，还用于：

对所述物理 I/O 设备进行模拟，以产生所述物理 I/O 设备对应的模拟设备；

25 将所述虚拟 I/O 设备关联的第二层队列绑定至所述模拟设备，以将所述虚拟 I/O 设备关联的第二层队列对接至所述物理 I/O 设备。

6、根据权利要求 4 所述的系统，所述队列组件在利用所述设备队列对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求进行调度的过程中，用于：

30 在所述实例向所述虚拟 I/O 设备发起读请求的过程中，读取所述实例下的各个 VCPU 的负载信息；

根据所述负载信息将所述读请求调度至所述虚拟 I/O 设备关联的第一层队列中的第一队列，以使用所述第一队列所对接的 VCPU 处理所述读请求。

35 7、根据权利要求 6 所述的系统，所述 I/O 卸载卡，还用于：获取所述读请求对应的响应报文；将所述响应报文中的元数据信息加入所述虚拟 I/O 设备关联的第二层队列中的指定队列；

所述队列组件在根据所述负载信息将所述读请求调度至所述虚拟 I/O 设备关联的第一层队列中的第一队列的过程中，用于：根据所述负载信息将所述指定队列中的所述元数据信息调度至所述第一队列中。

- 5 8、根据权利要求 7 所述的系统，所述 I/O 卸载卡在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据的过程中，用于：将所述响应报文中的数据部分写入所述实例对应的内存空间；将所述数据部分所在的内存地址添加至所述元数据信息中；

所述队列组件，还用于：触发目标队列对接的 VCPU 按照所述元数据信息中的内存地址读取所述响应报文中的数据部分。

- 10 9、根据权利要求 4 所述的系统，所述队列组件在利用所述设备队列对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求进行调度过程中，用于：将所述实例向所述虚拟 I/O 设备发起的写请求的元数据信息加入所述虚拟 I/O 设备关联的第一层队列中；将所述写请求的元数据信息调度至所述虚拟 I/O 设备关联的第二层队列中的第二队列；

- 15 所述 I/O 卸载卡，还用于：利用所述第二队列对接的物理 I/O 设备处理所述写请求。

10、根据权利要求 9 所述的系统，所述队列组件，还用于：将所述写请求对应的数据部分写入所述实例对应的内存空间；将所述数据部分所在的内存地址添加至所述元数据信息中；

- 20 所述 I/O 卸载卡在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据的过程中，用于：从所述第二队列中读取所述元数据信息；按照所述元数据信息中的内存地址获取所述写请求的数据部分；将所述数据部分发送至所述第二队列对接的所述物理 I/O 设备中。

11、根据权利要求 1 所述的系统，所述队列组件在为所述宿主机上的实例提供虚拟 I/O 设备的过程中，用于：

- 25 对需要与所述实例进行 I/O 的物理 I/O 设备进行 I/O 虚拟化，以产生所述物理 I/O 设备对应的虚拟 I/O 设备。

12、根据权利要求 11 所述的系统，所述队列组件在对需要与所述实例进行 I/O 的物理 I/O 设备进行 I/O 虚拟化的过程中，用于：

- 30 采用 SRIOV 技术，为需要与所述实例进行 I/O 的物理 I/O 设备创建 VF 设备，所述 VF 设备用于与所述实例下的各个 VCPU 进行数据交换；

利用所述实例的操作系统将所述 VF 设备注册为指定类型的虚拟 I/O 设备。

13、根据权利要求 1 所述的系统，所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求采用 io-uring 协议。

- 35 14、根据权利要求 1 所述的系统，所述队列组件中还包括加速模块，所述加速模块用于：

对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求所相关的数据执行加速处理，所述加速处理包括：加解密、压缩和统计卸载中的一种或多种处理。

15、根据权利要求 14 所述的系统，所述加速模块绑定至所述宿主机为所述虚拟 I/O 设备分配的内存地址，所述虚拟 I/O 设备将其与所述实例之间发生的 I/O 请求传递至所述设备队列；所述设备队列触发所述加速模块访问所述内存地址并对所述 I/O 请求所相关的数据进行加速处理；所述设备队列将所述 I/O 请求传递至所述 I/O 卸载卡，以使所述 I/O 卸载卡感知到所述 I/O 请求并从所述内存地址读取所述加速处理后的数据。

10 16、一种云环境下的 I/O 卸载方法，适用于宿主机中的 CPU，所述 CPU 中装配有队列组件，所述方法包括：

利用所述队列组件为所述宿主机上的实例提供虚拟 I/O 设备；

利用所述队列组件为所述虚拟 I/O 设备配置对应的设备队列；

15 在所述 CPU 内，利用所述设备队列对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求进行调度，以供所述宿主机上插接的 I/O 卸载卡监听所述设备队列中的 I/O 请求并在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

20 17、一种云环境下的 I/O 卸载方法，适用于宿主机上插接的 I/O 卸载卡，所述宿主机 CPU 中装配有队列组件，所述方法，包括：

从所述队列组件中的设备队列中监听所述宿主机上的实例与所述队列组件为所述实例提供的虚拟 I/O 设备之间发生的 I/O 请求；

获取监听到的 I/O 请求所相关的数据；

25 在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

18、一种处理器 CPU，安装在宿主机中，所述 CPU 内装配有队列组件，所述 CPU 用于执行一条或多条计算机指令，以用于：

利用所述队列组件为所述宿主机上的实例提供虚拟 I/O 设备；

30 利用所述队列组件为所述虚拟 I/O 设备配置对应的设备队列；

在所述 CPU 内，利用所述设备队列对所述实例与所述虚拟 I/O 设备之间发生的 I/O 请求进行调度，以供所述宿主机上插接的 I/O 卸载卡监听所述设备队列中的 I/O 请求并在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

35

19、一种 I/O 卸载卡，插接于宿主机上，所述宿主机的 CPU 内装配有队列组件，所述 I/O 卸载卡包括存储器和处理器；

所述存储器用于存储一条或多条计算机指令；

所述处理器与所述存储器耦合，用于执行所述一条或多条计算机指令，以用于：

5 从所述队列组件中的设备队列中监听所述宿主机上的实例与所述队列组件为所述实例提供的虚拟 I/O 设备之间发生的 I/O 请求；

获取监听到的 I/O 请求所相关的数据；

在所述实例与所述虚拟 I/O 设备对应的物理 I/O 设备之间传递监听到的 I/O 请求所相关的数据。

10

20、一种存储计算机指令的计算机可读存储介质，当所述计算机指令被一个或多个处理器执行时，致使所述一个或多个处理器执行权利要求 16 或 17 任一项所述的云环境下的 I/O 卸载方法。