



(12) 发明专利申请

(10) 申请公布号 CN 118819873 A

(43) 申请公布日 2024.10.22

(21) 申请号 202411318797.X

(22) 申请日 2024.09.21

(71) 申请人 珠海星云智联科技有限公司

地址 519000 广东省珠海市横琴新区宝华
路6号105室-74030

(72) 发明人 陈森法

(74) 专利代理机构 广州三环专利商标代理有限
公司 44202

专利代理师 陈舟苗

(51) Int. Cl.

G06F 9/50 (2006.01)

G06F 9/455 (2018.01)

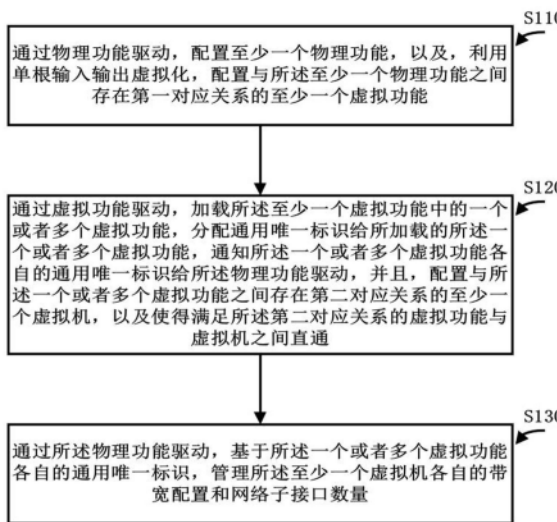
权利要求书3页 说明书14页 附图2页

(54) 发明名称

虚拟功能的管理方法、计算机设备、介质及系统

(57) 摘要

本申请涉及计算机技术领域并提供一种虚拟功能的管理方法、计算机设备、介质及系统。方法包括：通过物理功能驱动，配置至少一个物理功能以及配置与物理功能之间存在第一对应关系的至少一个虚拟功能；通过虚拟功能驱动，加载所述虚拟功能中的一个或者多个虚拟功能，分配通用唯一标识给所加载的虚拟功能，通知通用唯一标识给物理功能驱动，并且，配置与所述一个或者多个虚拟功能之间存在第二对应关系的至少一个虚拟机，以及使得满足第二对应关系的虚拟功能与虚拟机之间直通；通过物理功能驱动，基于通用唯一标识，管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。如此有利于高带宽利用率和高资源利用率。



1. 一种虚拟功能的管理方法,其特征在于,所述管理方法包括:

通过物理功能驱动,配置至少一个物理功能,以及,利用单根输入输出虚拟化,配置与所述至少一个物理功能之间存在第一对应关系的至少一个虚拟功能;

通过虚拟功能驱动,加载所述至少一个虚拟功能中的一个或者多个虚拟功能,分配通用唯一标识给所加载的所述一个或者多个虚拟功能,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,并且,配置与所述一个或者多个虚拟功能之间存在第二对应关系的至少一个虚拟机,以及使得满足所述第二对应关系的虚拟功能与虚拟机之间直通;

通过所述物理功能驱动,基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。

2. 根据权利要求1所述的管理方法,其特征在于,所述管理方法还包括:

通过所述虚拟功能驱动,移除所述至少一个虚拟功能中的第一虚拟功能,通知所述第一虚拟功能的通用唯一标识给所述物理功能驱动;

通过所述物理功能驱动,基于所述第一虚拟功能的通用唯一标识,确定所述至少一个虚拟机中与所述第一虚拟功能之间满足所述第二对应关系的第一虚拟机,删除所述第一虚拟机的关联所述第一虚拟功能的带宽配置,以及,回收所述第一虚拟机的关联所述第一虚拟功能的已分配网络子接口。

3. 根据权利要求1所述的管理方法,其特征在于,所述管理方法还包括:

通过所述物理功能驱动,基于所述一个或者多个虚拟功能各自的通用唯一标识,管理来自所述一个或者多个虚拟功能的资源申请。

4. 根据权利要求1所述的管理方法,其特征在于,所述物理功能驱动通过扩展网络接口配置命令从而实现:基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。

5. 根据权利要求1所述的管理方法,其特征在于,所述带宽配置包括下行最大带宽、上行最大带宽、源机器地址防欺诈检查。

6. 根据权利要求1所述的管理方法,其特征在于,所述物理功能驱动被部署在宿主机的内核态,所述虚拟功能驱动由系统模拟器提供,并且,所述系统模拟器被部署在所述宿主机的用户态,所述系统模拟器用于:解析所述一个或者多个虚拟功能各自的通用唯一标识从而确定与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,以及,通知与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识给所述物理功能驱动。

7. 根据权利要求6所述的管理方法,其特征在于,与所述宿主机连接的硬件侧包括微处理器,所述微处理器用于提供组策略管理装置,所述组策略管理装置用于为所述至少一个物理功能提供组级别的物理功能的发送速率限速配置和设备规格配置,所述组级别的物理功能包括所述至少一个物理功能中的至少两个物理功能。

8. 根据权利要求7所述的管理方法,其特征在于,所述组策略管理装置用于提供跨物理功能的硬件资源管理。

9. 根据权利要求1所述的管理方法,其特征在于,所述至少一个物理功能包括第一物理

功能和第二物理功能,所述至少一个虚拟功能中的第一虚拟功能集合与所述第一物理功能之间满足所述第一对应关系,所述至少一个虚拟功能中的第二虚拟功能集合与所述第二物理功能之间满足所述第一对应关系,所述至少一个虚拟机中的第二虚拟机与所述第一虚拟功能集合之间满足所述第二对应关系,所述第二虚拟机与所述第二虚拟功能集合之间满足所述第二对应关系,所述管理方法还包括:

通过组策略管理装置,对与所述第二虚拟机之间满足所述第二对应关系的虚拟功能,执行统一的发送速率限速配置和设备规格配置。

10. 根据权利要求9所述的管理方法,其特征在于,满足所述第一对应关系的物理功能和虚拟功能对应同一个以太网接口并且使用所述同一个以太网接口的带宽。

11. 根据权利要求10所述的管理方法,其特征在于,所述第二虚拟机使用所述第一物理功能和所述第二物理功能各自对应的以太网接口的带宽。

12. 一种计算机设备,其特征在于,所述计算机设备包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现根据权利要求1至11中任一项所述的方法。

13. 一种计算机可读存储介质,其特征在于,所述计算机可读存储介质存储有计算机指令,当所述计算机指令在计算机设备上运行时使得所述计算机设备执行根据权利要求1至11中任一项所述的方法。

14. 一种虚拟功能的管理系统,其特征在于,所述管理系统包括:

物理功能驱动,用于:配置至少一个物理功能,以及,利用单根输入输出虚拟化,配置与所述至少一个物理功能之间存在第一对应关系的至少一个虚拟功能;

虚拟功能驱动,用于:加载所述至少一个虚拟功能中的一个或者多个虚拟功能,分配通用唯一标识给所加载的所述一个或者多个虚拟功能,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,并且,配置与所述一个或者多个虚拟功能之间存在第二对应关系的至少一个虚拟机,以及使得满足所述第二对应关系的虚拟功能与虚拟机之间直通,

其中,所述物理功能驱动还用于:基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。

15. 根据权利要求14所述的管理系统,其特征在于,所述虚拟功能驱动还用于:移除所述至少一个虚拟功能中的第一虚拟功能,通知所述第一虚拟功能的通用唯一标识给所述物理功能驱动;

所述物理功能驱动还用于:基于所述第一虚拟功能的通用唯一标识,确定所述至少一个虚拟机中与所述第一虚拟功能之间满足所述第二对应关系的第一虚拟机,删除所述第一虚拟机的关联所述第一虚拟功能的带宽配置,以及,回收所述第一虚拟机的关联所述第一虚拟功能的已分配网络子接口。

16. 根据权利要求14所述的管理系统,其特征在于,所述物理功能驱动被部署在宿主机的内核态,所述虚拟功能驱动由系统模拟器提供,并且,所述系统模拟器被部署在所述宿主机的用户态,所述系统模拟器用于:解析所述一个或者多个虚拟功能各自的通用唯一标识从而确定与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,

以及,通知与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识给所述物理功能驱动。

17.根据权利要求16所述的管理系统,其特征在于,与所述宿主机连接的硬件侧包括微处理器,所述微处理器用于提供组策略管理装置,所述组策略管理装置用于为所述至少一个物理功能提供组级别的物理功能的发送速率限速配置和设备规格配置,所述组级别的物理功能包括所述至少一个物理功能中的至少两个物理功能。

18.根据权利要求14所述的管理系统,其特征在于,所述至少一个物理功能包括第一物理功能和第二物理功能,所述至少一个虚拟功能中的第一虚拟功能集合与所述第一物理功能之间满足所述第一对应关系,所述至少一个虚拟功能中的第二虚拟功能集合与所述第二物理功能之间满足所述第一对应关系,所述至少一个虚拟机中的第二虚拟机与所述第一虚拟功能集合之间满足所述第二对应关系,所述第二虚拟机与所述第二虚拟功能集合之间满足所述第二对应关系,所述管理系统还包括组策略管理装置,

所述组策略管理装置用于:对与所述第二虚拟机之间满足所述第二对应关系的虚拟功能,执行统一的发送速率限速配置和设备规格配置。

虚拟功能的管理方法、计算机设备、介质及系统

技术领域

[0001] 本申请涉及计算机技术领域,尤其涉及一种虚拟功能的管理方法、计算机设备、介质及系统。

背景技术

[0002] 在云计算、大型数据中心等应用中,将输入输出资源虚拟化后分配给各个虚拟机。利用单根输入输出虚拟化(Single Root I/O Virtualization,SRIOV),将一个物理功能虚拟出多个虚拟功能,再将这些虚拟功能直通给虚拟机,从而降低时延和提高吞吐率。但是,对于同一个虚拟机的用户而言,可能需要多个网口或者多个存储盘,因此一个虚拟机可能直通多个虚拟功能,而这些虚拟功能之间可能没有关联关系,因此难以对整个虚拟机的多个虚拟功能进行整体管理。现有技术中,一种管理方式是对每个虚拟功能单独限速,但是这样导致带宽利用率低下,另一种管理方式是对虚拟机中网络子接口数量统一限制,但是这样要求每个虚拟功能按照统一限制或者均分网络子接口数量,也是导致资源利用率低下。

[0003] 为此,本申请提供了一种虚拟功能的管理方法、计算机设备、介质及系统,用于应对现有技术中的技术难题。

发明内容

[0004] 第一方面,本申请提供了一种虚拟功能的管理方法。所述管理方法包括:通过物理功能驱动,配置至少一个物理功能,以及,利用单根输入输出虚拟化,配置与所述至少一个物理功能之间存在第一对应关系的至少一个虚拟功能;通过虚拟功能驱动,加载所述至少一个虚拟功能中的一个或者多个虚拟功能,分配通用唯一标识给所加载的所述一个或者多个虚拟功能,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,并且,配置与所述一个或者多个虚拟功能之间存在第二对应关系的至少一个虚拟机,以及使得满足所述第二对应关系的虚拟功能与虚拟机之间直通;通过所述物理功能驱动,基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。

[0005] 通过本申请的第一方面,在构建第一对应关系和构建第二对应关系的基础上,分别确定了被配置的物理功能与被配置的虚拟功能之间的第一对应关系以及被加载的虚拟功能与虚拟机之间的第二对应关系,并且,对虚拟功能的加载过程进行了优化,虚拟功能驱动在加载虚拟功能时分配了通用唯一标识,并且,还通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,这样物理功能驱动可以利用通用唯一标识来区分每个被加载的虚拟功能;一方面保持了虚拟机与被加载的虚拟功能之间直通,另一方面,只有被虚拟功能驱动加载的虚拟功能,其通用唯一标识才会被通知到物理功能驱动,这意味着,只有实际上被利用的虚拟功能通道才会被纳入物理功能驱动的管理范围,因此有效地克服了对每个虚拟功能单独设定最大数据传输速率或者最大网络子接口数量所导致的问题;并且,因为物理功能驱动可以利用通用唯一标识来精细化到每个物理功能各自对应的虚拟功

能,也就是精细化地管理各个物理功能对应的虚拟功能中被加载且被分配了通用唯一标识的虚拟功能,因此可以在整体资源管理策略的框架内进行灵活丰富的配置;因为是通过物理功能驱动来提供额外的资源管理接口,而虚拟功能是通过单根输入输出虚拟化从物理功能虚拟化得到,所以通过物理功能驱动来管理虚拟机各自的带宽配置和网络子接口数量,不会影响数据面的直通,同时又保持了控制面的灵活性;实现了高带宽利用率高资源利用率,以及可以灵活地应对虚拟机的虚拟功能资源管理方面的复杂性。

[0006] 在本申请的第一方面的一种可能的实现方式中,所述管理方法还包括:通过所述虚拟功能驱动,移除所述至少一个虚拟功能中的第一虚拟功能,通知所述第一虚拟功能的通用唯一标识给所述物理功能驱动;通过所述物理功能驱动,基于所述第一虚拟功能的通用唯一标识,确定所述至少一个虚拟机中与所述第一虚拟功能之间满足所述第二对应关系的第一虚拟机,删除所述第一虚拟机的关联所述第一虚拟功能的带宽配置,以及,回收所述第一虚拟机的关联所述第一虚拟功能的已分配网络子接口。

[0007] 在本申请的第一方面的一种可能的实现方式中,所述管理方法还包括:通过所述物理功能驱动,基于所述一个或者多个虚拟功能各自的通用唯一标识,管理来自所述一个或者多个虚拟功能的资源申请。

[0008] 在本申请的第一方面的一种可能的实现方式中,所述物理功能驱动通过扩展网络接口配置命令从而实现:基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。

[0009] 在本申请的第一方面的一种可能的实现方式中,所述带宽配置包括下行最大带宽、上行最大带宽、源机器地址防欺诈检查。

[0010] 在本申请的第一方面的一种可能的实现方式中,所述物理功能驱动被部署在宿主机的内核态,所述虚拟功能驱动由系统模拟器提供,并且,所述系统模拟器被部署在所述宿主机的用户态,所述系统模拟器用于:解析所述一个或者多个虚拟功能各自的通用唯一标识从而确定与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,以及,通知与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识给所述物理功能驱动。

[0011] 在本申请的第一方面的一种可能的实现方式中,与所述宿主机连接的硬件侧包括微处理器,所述微处理器用于提供组策略管理装置,所述组策略管理装置用于为所述至少一个物理功能提供组级别的物理功能的发送速率限速配置和设备规格配置,所述组级别的物理功能包括所述至少一个物理功能中的至少两个物理功能。

[0012] 在本申请的第一方面的一种可能的实现方式中,所述组策略管理装置用于提供跨物理功能的硬件资源管理。

[0013] 在本申请的第一方面的一种可能的实现方式中,所述至少一个物理功能包括第一物理功能和第二物理功能,所述至少一个虚拟功能中的第一虚拟功能集合与所述第一物理功能之间满足所述第一对应关系,所述至少一个虚拟功能中的第二虚拟功能集合与所述第二物理功能之间满足所述第一对应关系,所述至少一个虚拟机中的第二虚拟机与所述第一虚拟功能集合之间满足所述第二对应关系,所述第二虚拟机与所述第二虚拟功能集合之间满足所述第二对应关系,所述管理方法还包括:通过组策略管理装置,对与所述第二虚拟机

之间满足所述第二对应关系的虚拟功能,执行统一的发送速率限速配置和设备规格配置。

[0014] 在本申请的第一方面的一种可能的实现方式中,满足所述第一对应关系的物理功能和虚拟功能对应同一个以太网接口并且使用所述同一个以太网接口的带宽。

[0015] 在本申请的第一方面的一种可能的实现方式中,所述第二虚拟机使用所述第一物理功能和所述第二物理功能各自对应的以太网接口的带宽。

[0016] 第二方面,本申请实施例还提供了一种计算机设备,所述计算机设备包括存储器、处理器及存储在所述存储器上并可在所述处理器上运行的计算机程序,所述处理器执行所述计算机程序时实现根据上述任一方面的任一种实现方式的方法。

[0017] 第三方面,本申请实施例还提供了一种计算机可读存储介质,所述计算机可读存储介质存储有计算机指令,当所述计算机指令在计算机设备上运行时使得所述计算机设备执行根据上述任一方面的任一种实现方式的方法。

[0018] 第四方面,本申请实施例还提供了一种计算机程序产品,所述计算机程序产品包括存储在计算机可读存储介质上的指令,当所述指令在计算机设备上运行时使得所述计算机设备执行根据上述任一方面的任一种实现方式的方法。

[0019] 第五方面,本申请实施例还提供了一种虚拟功能的管理系统。所述管理系统包括:物理功能驱动,用于:配置至少一个物理功能,以及,利用单根输入输出虚拟化,配置与至少一个物理功能之间存在第一对应关系的至少一个虚拟功能;虚拟功能驱动,用于:加载所述至少一个虚拟功能中的一个或者多个虚拟功能,分配通用唯一标识给所加载的所述一个或者多个虚拟功能,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,并且,配置与所述一个或者多个虚拟功能之间存在第二对应关系的至少一个虚拟机,以及使得满足所述第二对应关系的虚拟功能与虚拟机之间直通,其中,所述物理功能驱动还用于:基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。

[0020] 通过本申请的第五方面,在构建第一对应关系和构建第二对应关系的基础上,分别确定了被配置的物理功能与被配置的虚拟功能之间的第一对应关系以及被加载的虚拟功能与虚拟机之间的第二对应关系,并且,对虚拟功能的加载过程进行了优化,虚拟功能驱动在加载虚拟功能时分配了通用唯一标识,并且,还通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,这样物理功能驱动可以利用通用唯一标识来区分每个被加载的虚拟功能;一方面保持了虚拟机与被加载的虚拟功能之间直通,另一方面,只有被虚拟功能驱动加载的虚拟功能,其通用唯一标识才会被通知到物理功能驱动,这意味着,只有实际上被利用的虚拟功能通道才会被纳入物理功能驱动的管理范围,因此有效地克服了对每个虚拟功能单独设定最大数据传输速率或者最大网络子接口数量所导致的问题;并且,因为物理功能驱动可以利用通用唯一标识来精细化到每个物理功能各自对应的虚拟功能,也就是精细化地管理各个物理功能对应的虚拟功能中被加载且被分配了通用唯一标识的虚拟功能,因此可以在整体资源管理策略的框架内进行灵活丰富的配置;因为是通过物理功能驱动来提供额外的资源管理接口,而虚拟功能是通过单根输入输出虚拟化从物理功能虚拟化得到,所以通过物理功能驱动来管理虚拟机各自的带宽配置和网络子接口数量,不会影响数据面的直通,同时又保持了控制面的灵活性;实现了高带宽利用率和高资源利用率,以及可以灵活地应对虚拟机的虚拟功能资源管理方面的复杂性。

[0021] 在本申请的第五方面的一种可能的实现方式中,所述虚拟功能驱动还用于:移除所述至少一个虚拟功能中的第一虚拟功能,通知所述第一虚拟功能的通用唯一标识给所述物理功能驱动;所述物理功能驱动还用于:基于所述第一虚拟功能的通用唯一标识,确定所述至少一个虚拟机中与所述第一虚拟功能之间满足所述第二对应关系的第一虚拟机,删除所述第一虚拟机的关联所述第一虚拟功能的带宽配置,以及,回收所述第一虚拟机的关联所述第一虚拟功能的已分配网络子接口。

[0022] 在本申请的第五方面的一种可能的实现方式中,所述物理功能驱动被部署在宿主机的内核态,所述虚拟功能驱动由系统模拟器提供,并且,所述系统模拟器被部署在所述宿主机的用户态,所述系统模拟器用于:解析所述一个或者多个虚拟功能各自的通用唯一标识从而确定与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,以及,通知与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识给所述物理功能驱动。

[0023] 在本申请的第五方面的一种可能的实现方式中,与所述宿主机连接的硬件侧包括微处理器,所述微处理器用于提供组策略管理装置,所述组策略管理装置用于为所述至少一个物理功能提供组级别的物理功能的发送速率限速配置和设备规格配置,所述组级别的物理功能包括所述至少一个物理功能中的至少两个物理功能。

[0024] 在本申请的第五方面的一种可能的实现方式中,所述至少一个物理功能包括第一物理功能和第二物理功能,所述至少一个虚拟功能中的第一虚拟功能集合与所述第一物理功能之间满足所述第一对应关系,所述至少一个虚拟功能中的第二虚拟功能集合与所述第二物理功能之间满足所述第一对应关系,所述至少一个虚拟机中的第二虚拟机与所述第一虚拟功能集合之间满足所述第二对应关系,所述第二虚拟机与所述第二虚拟功能集合之间满足所述第二对应关系,所述管理系统还包括组策略管理装置,所述组策略管理装置用于:对与所述第二虚拟机之间满足所述第二对应关系的虚拟功能,执行统一的发送速率限速配置和设备规格配置。

附图说明

[0025] 为了更清楚地说明本申请实施例技术方案,下面将对实施例描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本申请的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0026] 图1为本申请实施例提供的一种虚拟功能的管理方法的流程示意图;

图2为本申请实施例提供的一种虚拟功能的管理系统的示意图;

图3为本申请实施例提供的一种计算设备的结构示意图。

具体实施方式

[0027] 下面将结合附图对本申请实施例作进一步地详细描述。

[0028] 应当理解的是,在本申请的描述中,“至少一个”指一个或一个以上,“多个”指两个或两个以上。另外,“第一”、“第二”等词汇,除非另有说明,否则仅用于区分描述的目的,而不能理解为指示或暗示相对重要性,也不能理解为指示或暗示顺序。

[0029] 图1为本申请实施例提供的一种虚拟功能的管理方法的流程示意图。如图1所示,所述管理方法包括以下步骤。

[0030] 步骤S110:通过物理功能驱动,配置至少一个物理功能,以及,利用单根输入输出虚拟化,配置与所述至少一个物理功能之间存在第一对应关系的至少一个虚拟功能。

[0031] 步骤S120:通过虚拟功能驱动,加载所述至少一个虚拟功能中的一个或者多个虚拟功能,分配通用唯一标识给所加载的所述一个或者多个虚拟功能,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,并且,配置与所述一个或者多个虚拟功能之间存在第二对应关系的至少一个虚拟机,以及使得满足所述第二对应关系的虚拟功能与虚拟机之间直通。

[0032] 步骤S130:通过所述物理功能驱动,基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。

[0033] 参阅图1,图1所示的虚拟功能的管理方法,可以应用于如云计算、大型数据中心等技术领域,可以用于如云游戏、短视频、网络功能虚拟化(Network Function Virtualization,NFV)等,这些技术领域和应用对于数据吞吐率和网络时延有较高的要求。为了实现用户定制化输入输出资源以及为了提高资源利用率,利用单根输入输出虚拟化(Single Root I/O Virtualization,SRIOV),将一个物理功能虚拟出多个虚拟功能,再将这些虚拟功能直通给虚拟机,如此可以提高虚拟功能的输入输出处理性能,从而降低时延和提高吞吐率。一个虚拟机可能与多个虚拟功能之间建立起直通关系,该虚拟机可以通过多个虚拟功能实现多个网口的报文收发,多个存储盘的数据读写,或者多个数据通道等,这样该虚拟机的用户可以在定制化输入输出资源方面有更高的灵活性和操作空间。但是,与同一个虚拟机之间建立起直通关系的多个虚拟功能,可能来自于同一个物理功能,也可能来自于不同的物理功能,因此,与同一个虚拟机之间建立起直通关系的多个虚拟功能之间的关联关系是难以预计的,这样就给管理整个虚拟机的虚拟功能资源带来了挑战。如果对每个虚拟功能单独设定最大数据传输速率或者最大网络子接口数量,则不利于充分利用有限的带宽资源和接口资源,这样不利于云厂商管理虚拟机和用户使用虚拟机,而且如果每个虚拟按照单独限速方案来执行带宽限速,则意味着最大带宽是受到已启用的虚拟功能的数量的限制,只有当每个虚拟功能均参与到输入输出时才能利用所有带宽,这一点对于云厂商和用户都是不友好的。下面结合本申请具体实施例详细说明,图1所示的虚拟功能的管理方法,如何实现了高带宽利用率高资源利用率,以及可以灵活地应对虚拟机的虚拟功能资源管理方面的复杂性。

[0034] 继续参阅图1,在步骤S110,通过物理功能驱动,配置至少一个物理功能,以及,利用单根输入输出虚拟化,配置与所述至少一个物理功能之间存在第一对应关系的至少一个虚拟功能。如此,实现了灵活配置输入输出资源,第一对应关系定义了所述至少一个物理功能与所述至少一个虚拟功能之间的对应关系,因此,通过解析第一对应关系,可以确定某个虚拟功能所对应的物理功能,或者,可以确定利用单根输入输出虚拟化将某个物理功能虚拟得到的对应的虚拟功能。然后,在步骤S120,通过虚拟功能驱动,加载所述至少一个虚拟功能中的一个或者多个虚拟功能,分配通用唯一标识给所加载的所述一个或者多个虚拟功能,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,并且,配置与所述一个或者多个虚拟功能之间存在第二对应关系的至少一个虚拟机,以及使得满足所

述第二对应关系的虚拟功能与虚拟机之间直通。如此,在步骤S110构建的物理功能与虚拟功能之间的第一对应关系的基础上,在步骤S120,对虚拟功能的加载过程进行了优化。应当理解的是,加载所述至少一个虚拟功能中的一个或者多个虚拟功能,这意味着所述一个或者多个虚拟功能是被加载的虚拟功能也即实际上启用的虚拟功能,而没有被加载的虚拟功能则当前没有被启用。针对被加载的虚拟功能也即所述一个或者多个虚拟功能,虚拟功能驱动在加载虚拟功能时分配了通用唯一标识(Universally Unique Identifier,UUID)给虚拟功能,因此每个被加载也即被启用的虚拟功能均被分配了通用唯一标识用于区分每个被加载的虚拟功能。虚拟功能驱动还通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,这样物理功能驱动可以利用通用唯一标识来区分每个被加载的虚拟功能。虚拟功能驱动还配置与所述一个或者多个虚拟功能之间存在第二对应关系的至少一个虚拟机,以及使得满足所述第二对应关系的虚拟功能与虚拟机之间直通。如此,实现了被加载的所述一个或者多个虚拟功能与虚拟机之间的直通,可以提高虚拟功能的输入输出处理性能,从而降低时延和提高吞吐率。第二对应关系定义了被加载的虚拟功能与虚拟机之间的对应关系,因此,通过解析第二对应关系,可以确定某个虚拟功能所对应(也即存在直通关系)的虚拟机,或者,可以确定某个虚拟机所对应的所有虚拟功能(也即存在直通关系)。

[0035] 继续参阅图1,在步骤S110所构建的第一对应关系和在步骤S120所构建的第二对应关系的基础上,分别确定了被配置物理功能与被配置的虚拟功能之间的第一对应关系以及被加载的虚拟功能与虚拟机之间的第二对应关系,并且,对虚拟功能的加载过程进行了优化,虚拟功能驱动在加载虚拟功能时分配了通用唯一标识,并且,还通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,这样物理功能驱动可以利用通用唯一标识来区分每个被加载的虚拟功能。然后,在步骤S130,通过所述物理功能驱动,基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。如此,在虚拟机与被加载的虚拟功能之间直通的基础上,对虚拟功能的加载过程进行了优化,虚拟功能驱动为每个被加载的虚拟功能分配通用唯一标识,这样物理功能驱动可以基于通用唯一标识来区分和管理与虚拟功能对应的被加载的虚拟功能,并且,物理功能驱动可以提供基于通用唯一标识的资源管理接口,例如,配置下行最大带宽,配置上行最大带宽,配置子接口数量,配置源物理地址防欺诈检查等,还可以通过通用唯一标识为虚拟机创建私有配置和私有配置文件还有子目录。因为是通过物理功能驱动来管理所述至少一个虚拟机各自的带宽配置和网络子接口数量,因此可以精细化到每个物理功能各自对应的虚拟功能,而利用所述一个或者多个虚拟功能各自的通用唯一标识,可以针对被加载的也即已被启用的虚拟功能进行如带宽限速和接口数量限制。因此,一方面保持了虚拟机与被加载的虚拟功能之间直通,另一方面,只有被虚拟功能驱动加载的虚拟功能,其通用唯一标识才会被通知到物理功能驱动,这意味着,只有实际上被利用的虚拟功能通道才会被纳入物理功能驱动的管理范围,因此有效地克服了对每个虚拟功能单独设定最大数据传输速率或者最大网络子接口数量所导致的问题。并且,因为物理功能驱动可以利用通用唯一标识来精细化到每个物理功能各自对应的虚拟功能,也就是精细化地管理各个物理功能对应的虚拟功能中被加载且被分配了通用唯一标识的虚拟功能,因此可以在整体资源管理策略的框架内进行灵活丰富的配置,例如,配置下行最大带宽,配置上行最大带

宽,配置子接口数量,配置源物理地址防欺诈检查等,还可以通过通用唯一标识为虚拟机创建私有配置和私有配置文件还有子目录,从而更好满足用户定制化需求。进一步地,因为是通过物理功能驱动来提供额外的资源管理接口,而虚拟功能是通过单根输入输出虚拟化从物理功能虚拟化得到,所以通过物理功能驱动来管理虚拟机各自的带宽配置和网络子接口数量,不会影响数据面的直通,同时又保持了控制面的灵活性。进一步地,图1所示的虚拟功能的管理方法,对虚拟功能的加载过程进行了优化,虚拟功能驱动在加载虚拟功能时分配了通用唯一标识,并且,还通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,这样物理功能驱动可以利用通用唯一标识来区分每个被加载的虚拟功能;如此,便于应用到既有的物理功能与虚拟功能的配置方案,并且,对内核态的物理功能驱动做出的改动幅度很小,可以通过扩展物理功能驱动网络接口配置命令(IP LINK),例如通过扩展命令行命令和增加参数化配置,不仅可以实现整体资源规划,也可以实现在定制化输入输出资源方面有更高的灵活性和操作空间。

[0036] 总之,图1所示的虚拟功能的管理方法,在构建第一对应关系和构建第二对应关系的基础上,分别确定了被配置的物理功能与被配置的虚拟功能之间的第一对应关系以及被加载的虚拟功能与虚拟机之间的第二对应关系,并且,对虚拟功能的加载过程进行了优化,虚拟功能驱动在加载虚拟功能时分配了通用唯一标识,并且,还通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,这样物理功能驱动可以利用通用唯一标识来区分每个被加载的虚拟功能;一方面保持了虚拟机与被加载的虚拟功能之间直通,另一方面,只有被虚拟功能驱动加载的虚拟功能,其通用唯一标识才会被通知到物理功能驱动,这意味着,只有实际上被利用的虚拟功能通道才会被纳入物理功能驱动的管理范围,因此有效地克服了对每个虚拟功能单独设定最大数据传输速率或者最大网络子接口数量所导致的问题;并且,因为物理功能驱动可以利用通用唯一标识来精细化到每个物理功能各自对应的虚拟功能,也就是精细化地管理各个物理功能对应的虚拟功能中被加载且被分配了通用唯一标识的虚拟功能,因此可以在整体资源管理策略的框架内进行灵活丰富的配置;因为是通过物理功能驱动来提供额外的资源管理接口,而虚拟功能是通过单根输入输出虚拟化从物理功能虚拟化得到,所以通过物理功能驱动来管理虚拟机各自的带宽配置和网络子接口数量,不会影响数据面的直通,同时又保持了控制面的灵活性;实现了高带宽利用率和资源利用率,以及可以灵活地应对虚拟机的虚拟功能资源管理方面的复杂性。

[0037] 在一种可能的实现方式中,所述管理方法还包括:通过所述虚拟功能驱动,移除所述至少一个虚拟功能中的第一虚拟功能,通知所述第一虚拟功能的通用唯一标识给所述物理功能驱动;通过所述物理功能驱动,基于所述第一虚拟功能的通用唯一标识,确定所述至少一个虚拟机中与所述第一虚拟功能之间满足所述第二对应关系的第一虚拟机,删除所述第一虚拟机的关联所述第一虚拟功能的带宽配置,以及,回收所述第一虚拟机的关联所述第一虚拟功能的已分配网络子接口。如此,物理功能驱动可以利用通用唯一标识来精细化到每个物理功能各自对应的虚拟功能,也就是精细化地管理各个物理功能对应的虚拟功能中被加载且被分配了通用唯一标识的虚拟功能,因此可以在整体资源管理策略的框架内进行灵活丰富的配置,提供了虚拟功能移除和相应的带宽配置删除以及已分配网络子接口的回收,有利于资源回收再利用和提高资源利用率。

[0038] 在一种可能的实现方式中,所述管理方法还包括:通过所述物理功能驱动,基于所

述一个或者多个虚拟功能各自的通用唯一标识,管理来自所述一个或者多个虚拟功能的资源申请。如此,物理功能驱动可以利用通用唯一标识来精细化到每个物理功能各自对应的虚拟功能,也就是精细化地管理各个物理功能对应的虚拟功能中被加载且被分配了通用唯一标识的虚拟功能,因此可以在整体资源管理策略的框架内进行灵活丰富的配置。

[0039] 在一种可能的实现方式中,所述物理功能驱动通过扩展网络接口配置命令从而实现:基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。如此,便于应用到既有的物理功能与虚拟功能的配置方案,并且,对内核态的物理功能驱动做出的改动幅度很小,可以通过扩展物理功能驱动网络接口配置命令(IP LINK),例如通过扩展命令行命令和增加参数化配置,不仅可以实现整体资源规划,也可以实现在定制化输入输出资源方面有更高的灵活性和操作空间。

[0040] 在一种可能的实现方式中,所述带宽配置包括下行最大带宽、上行最大带宽、源机器地址防欺诈检查。如此,物理功能驱动可以利用通用唯一标识来精细化到每个物理功能各自对应的虚拟功能,也就是精细化地管理各个物理功能对应的虚拟功能中被加载且被分配了通用唯一标识的虚拟功能,因此可以在整体资源管理策略的框架内进行灵活丰富的配置。

[0041] 在一种可能的实现方式中,所述物理功能驱动被部署在宿主机的内核态,所述虚拟功能驱动由系统模拟器提供,并且,所述系统模拟器被部署在所述宿主机的用户态,所述系统模拟器用于:解析所述一个或者多个虚拟功能各自的通用唯一标识从而确定与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动,以及,通知与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识给所述物理功能驱动。如此,可以适配各种虚拟化技术和产品,例如在云计算和云服务的应用中采用的开源的虚拟机监视器和模拟器,可用于模拟各种硬件平台。

[0042] 在一些实施例,与所述宿主机连接的硬件侧包括微处理器,所述微处理器用于提供组策略管理装置,所述组策略管理装置用于为所述至少一个物理功能提供组级别的物理功能的发送速率限速配置和设备规格配置,所述组级别的物理功能包括所述至少一个物理功能中的至少两个物理功能。如此,在上述的通过所述物理功能驱动实现的管理基础上(通过所述物理功能驱动,基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量),提供了额外的管理手段,因此,物理功能驱动维护自身虚拟功能与通用唯一标识的管理,在硬件侧增加了微处理器,例如微控制单元(Microcontroller Unit,MCU),这样可以支持多个物理功能各自对应的虚拟功能加入同一个虚拟机。如此,通过支持多个物理功能各自对应的虚拟功能加入同一个虚拟机,以及为这些虚拟功能的资源管理提供组策略管理装置,有助于充分利用硬件的输入输出带宽和提高系统可靠性,也有利于云厂商基于整个硬件资源来进行管理,而不受限于物理功能对硬件资源的划分。例如,以网卡为例,一个物理功能以及其对应的虚拟功能对应一个以太网面板口,通过支持多个物理功能各自对应的虚拟功能加入同一个虚拟机,可以使用多个以太网面板口的带宽,这样可以避免因为一个以太网面板口故障而导致整个虚拟机的流量中断。在一些实施例,所述组策略管理装置用于提供跨物理功能的硬件资源管理。如此,有利于云厂商基于整个硬件资源来进行管理,而不受限于物理功能对硬件资源的划分。

[0043] 在一种可能的实现方式中,所述至少一个物理功能包括第一物理功能和第二物理功能,所述至少一个虚拟功能中的第一虚拟功能集合与所述第一物理功能之间满足所述第一对应关系,所述至少一个虚拟功能中的第二虚拟功能集合与所述第二物理功能之间满足所述第二对应关系,所述至少一个虚拟机中的第二虚拟机与所述第一虚拟功能集合之间满足所述第二对应关系,所述第二虚拟机与所述第二虚拟功能集合之间满足所述第二对应关系,所述管理方法还包括:通过组策略管理装置,对与所述第二虚拟机之间满足所述第二对应关系的虚拟功能,执行统一的发送速率限速配置和设备规格配置。如此,有利于云厂商基于整个硬件资源来进行管理,而不受限于物理功能对硬件资源的划分。

[0044] 在一些实施例中,满足所述第一对应关系的物理功能和虚拟功能对应同一个以太网接口并且使用所述同一个以太网接口的带宽。在一些实施例中,所述第二虚拟机使用所述第一物理功能和所述第二物理功能各自对应的以太网接口的带宽。如此,通过支持多个物理功能各自对应的虚拟功能加入同一个虚拟机,以及为这些虚拟功能的资源管理提供组策略管理装置,有助于充分利用硬件的输入输出带宽和提高系统可靠性,可以使用多个以太网接口的带宽,这样可以避免因为一个以太网接口故障而导致整个虚拟机的流量中断。

[0045] 图2为本申请实施例提供的一种虚拟功能的管理系统的示意图。如图2所示,所述管理系统包括:物理功能驱动202,用于:配置至少一个物理功能,以及,利用单根输入输出虚拟化,配置与所述至少一个物理功能之间存在第一对应关系的至少一个虚拟功能;虚拟功能驱动204,用于:加载所述至少一个虚拟功能中的一个或者多个虚拟功能,分配通用唯一标识给所加载的所述一个或者多个虚拟功能,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动202,并且,配置与所述一个或者多个虚拟功能之间存在第二对应关系的至少一个虚拟机,以及使得满足所述第二对应关系的虚拟功能与虚拟机之间直通。其中,所述物理功能驱动202还用于:基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量。如图2所示,所述物理功能驱动202被部署在宿主机内核态250,所述虚拟功能驱动204由系统模拟器260提供,并且,所述系统模拟器260被部署在所述宿主机用户态252。所述系统模拟器260用于:解析所述一个或者多个虚拟功能各自的通用唯一标识从而确定与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识,通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动202,以及,通知与所述一个或者多个虚拟功能之间满足所述第一对应关系的物理功能的总线设备功能标识给所述物理功能驱动202。

[0046] 参阅图2,图2中示例性示出了至少一个物理功能和至少一个虚拟功能。第一对应关系定义了所述至少一个物理功能与所述至少一个虚拟功能之间的对应关系,因此,通过解析第一对应关系,可以确定某个虚拟功能所对应的物理功能,或者,可以确定利用单根输入输出虚拟化将某个物理功能虚拟得到的对应的虚拟功能。如图2所示,物理功能驱动202配置了物理功能A210和物理功能B220,还配置了物理功能A210的虚拟功能A212、物理功能A210的虚拟功能B214、物理功能B220的虚拟功能C222以及物理功能B220的虚拟功能D224。系统模拟器260可以是在云计算和云服务的应用中采用的开源的虚拟机监视器和模拟器(Quick EMulator, QEMU),可用于模拟多个硬件平台。一个虚拟机可能与多个虚拟功能之间建立起直通关系,该虚拟机可以通过多个虚拟功能实现多个网口的报文收发,多个存储盘

的数据读写,或者多个数据通道等,这样该虚拟机的用户可以在定制化输入输出资源方面有更高的灵活性和操作空间。但是,与同一个虚拟机之间建立起直通关系的多个虚拟功能,可能来自于同一个物理功能,也可能来自于不同的物理功能,因此,与同一个虚拟机之间建立起直通关系的多个虚拟功能之间的关联关系是难以预计的。例如,图2所示的物理功能A210的虚拟功能A212和物理功能B220的虚拟功能C222分别对应物理功能A210和物理功能B220,同一个虚拟机可能分别使用物理功能A210的虚拟功能A212和物理功能B220的虚拟功能C222来支持两个网口,或者用于支持两个存储盘的数据通道。因为物理功能驱动202可以利用通用唯一标识来精细化到每个物理功能各自对应的虚拟功能,也就是精细化地管理各个物理功能对应的虚拟功能中被加载且被分配了通用唯一标识的虚拟功能,因此可以在整体资源管理策略的框架内进行灵活丰富的配置。

[0047] 继续参阅图2,在一些实施例中,与所述宿主机连接的硬件侧254包括微处理器270,所述微处理器270用于提供组策略管理装置272,所述组策略管理装置272用于为所述至少一个物理功能提供组级别的物理功能的发送速率限速配置和设备规格配置,所述组级别的物理功能包括所述至少一个物理功能中的至少两个物理功能。如此,在上述的通过所述物理功能驱动202实现的管理基础上(通过所述物理功能驱动202,基于所述一个或者多个虚拟功能各自的通用唯一标识,管理所述至少一个虚拟机各自的带宽配置和网络子接口数量),提供了额外的管理手段,因此,物理功能驱动202维护自身虚拟功能与通用唯一标识的管理,在硬件侧254增加了微处理器270,例如微控制单元(Microcontroller Unit,MCU),这样可以支持多个物理功能各自对应的虚拟功能加入同一个虚拟机。如此,通过支持多个物理功能各自对应的虚拟功能加入同一个虚拟机,以及为这些虚拟功能的资源管理提供组策略管理装置272,有助于充分利用硬件的输入输出带宽和提高系统可靠性,也有利于云厂商基于整个硬件资源来进行管理,而不受限于物理功能对硬件资源的划分。例如,以网卡为例,一个物理功能以及其对应的虚拟功能对应一个以太网面板口,通过支持多个物理功能各自对应的虚拟功能加入同一个虚拟机,可以使用多个以太网面板口的带宽,这样可以避免因为一个以太网面板口故障而导致整个虚拟机的流量中断。

[0048] 总之,图2所示的虚拟功能的管理系统,在构建第一对应关系和构建第二对应关系的基础上,分别确定了被配置的物理功能与被配置的虚拟功能之间的第一对应关系以及被加载的虚拟功能与虚拟机之间的第二对应关系,并且,对虚拟功能的加载过程进行了优化,虚拟功能驱动204在加载虚拟功能时分配了通用唯一标识,并且,还通知所述一个或者多个虚拟功能各自的通用唯一标识给所述物理功能驱动202,这样物理功能驱动202可以利用通用唯一标识来区分每个被加载的虚拟功能;一方面保持了虚拟机与被加载的虚拟功能之间直通,另一方面,只有被虚拟功能驱动204加载的虚拟功能,其通用唯一标识才会被通知到物理功能驱动202,这意味着,只有实际上被利用的虚拟功能通道才会被纳入物理功能驱动202的管理范围,因此有效地克服了对每个虚拟功能单独设定最大数据传输速率或者最大网络子接口数量所导致的问题;并且,因为物理功能驱动202可以利用通用唯一标识来精细化到每个物理功能各自对应的虚拟功能,也就是精细化地管理各个物理功能对应的虚拟功能中被加载且被分配了通用唯一标识的虚拟功能,因此可以在整体资源管理策略的框架内进行灵活丰富的配置;因为是通过物理功能驱动202来提供额外的资源管理接口,而虚拟功能是通过单根输入输出虚拟化从物理功能虚拟化得到,所以通过物理功能驱动202来管理

虚拟机各自的带宽配置和网络子接口数量,不会影响数据面的直通,同时又保持了控制面的灵活性;实现了高带宽利用率高资源利用率,以及可以灵活地应对虚拟机的虚拟功能资源管理方面的复杂性。

[0049] 参阅图2,在一种可能的实施方式中,所述虚拟功能驱动204还用于:移除所述至少一个虚拟功能中的第一虚拟功能,通知所述第一虚拟功能的通用唯一标识给所述物理功能驱动202;所述物理功能驱动202还用于:基于所述第一虚拟功能的通用唯一标识,确定所述至少一个虚拟机中与所述第一虚拟功能之间满足所述第二对应关系的第一虚拟机,删除所述第一虚拟机的关联所述第一虚拟功能的带宽配置,以及,回收所述第一虚拟机的关联所述第一虚拟功能的已分配网络子接口。如此,提供了虚拟功能移除和相应的带宽配置删除以及已分配网络子接口的回收,有利于资源回收再利用和提高资源利用率。

[0050] 在一种可能的实施方式中,所述至少一个物理功能包括第一物理功能和第二物理功能,所述至少一个虚拟功能中的第一虚拟功能集合与所述第一物理功能之间满足所述第一对应关系,所述至少一个虚拟功能中的第二虚拟功能集合与所述第二物理功能之间满足所述第一对应关系,所述至少一个虚拟机中的第二虚拟机与所述第一虚拟功能集合之间满足所述第二对应关系,所述第二虚拟机与所述第二虚拟功能集合之间满足所述第二对应关系,所述管理系统还包括组策略管理装置272,所述组策略管理装置272用于:对与所述第二虚拟机之间满足所述第二对应关系的虚拟功能,执行统一的发送速率限速配置和设备规格配置。如此,有利于云厂商基于整个硬件资源来进行管理,而不受限于物理功能对硬件资源的划分。

[0051] 图3是本申请实施例提供的一种计算设备的结构示意图,该计算设备300包括:一个或者多个处理器310、通信接口320以及存储器330。所述处理器310、通信接口320以及存储器330通过总线340相互连接。可选地,该计算设备300还可以包括输入/输出接口350,输入/输出接口350连接有输入/输出设备,用于接收用户设置的参数等。该计算设备300能够用于实现上述的本申请实施例中设备实施例或者系统实施例的部分或者全部功能;处理器310还能够用于实现上述的本申请实施例中方法实施例的部分或者全部操作步骤。例如,该计算设备300执行各种操作的具体实现可参照上述实施例中的具体细节,如处理器310用于执行上述方法实施例中部分或者全部步骤或者上述方法实施例中的部分或者全部操作。再例如,本申请实施例中,计算设备300可用于实现上述装置实施例中一个或者多个部件的部分或者全部功能,此外通信接口320具体可用于为了实现这些装置、部件的功能所必须的通讯功能等,以及处理器310具体可用于为了实现这些装置、部件的功能所必须的处理功能等。

[0052] 应当理解的是,图3的计算设备300可以包括一个或者多个处理器310,并且多个处理器310可以按照并行化连接方式、串行化连接方式、串并行连接方式或者任意连接方式来协同提供处理能力,或者多个处理器310可以构成处理器序列或者处理器阵列,或者多个处理器310之间可以分成主处理器和辅助处理器,或者多个处理器310之间可以具有不同的架构如采用异构计算架构。另外,图3所示的计算设备300,相关的结构性描述及功能性描述是示例性且非限制性的。在一些示例性实施例中,计算设备300可以包括比图3所示的更多或更少的部件,或者组合某些部件,或者拆分某些部件,或者具有不同的部件布置。

[0053] 处理器310可以有多种具体实现形式,例如处理器310可以包括中央处理器

(central processing unit,CPU)、图形处理器(graphic processing unit,GPU)、神经网络处理器(neural-network processing unit,NPU)、张量处理器(tensor processing unit,TPU)或数据处理器(data processing unit,DPU)等一种或多种的组合,本申请实施例不做具体限定。处理器310还可以是单核处理器或多核处理器。处理器310可以由CPU和硬件芯片的组合。上述硬件芯片可以是专用集成电路(application-specific integrated circuit,ASIC),可编程逻辑器件(programmable logic device,PLD)或其组合。上述PLD可以是复杂可编程逻辑器件(complex programmable logic device,CPLD),现场可编程逻辑门阵列(field-programmable gate array,FPGA),通用阵列逻辑(generic array logic,GAL)或其任意组合。处理器310也可以单独采用内置处理逻辑的逻辑器件来实现,例如FPGA或数字信号处理器(digital signal processor,DSP)等。通信接口320可以为有线接口或无线接口,用于与其他模块或设备进行通信,有线接口可以是以太接口、局域互连网络(local interconnect network,LIN)等,无线接口可以是蜂窝网络接口或使用无线局域网接口等。

[0054] 存储器330可以是非易失性存储器,例如,只读存储器(read-only memory,ROM)、可编程只读存储器(programmable ROM,PROM)、可擦除可编程只读存储器(erasable PROM,EPROM)、电可擦除可编程只读存储器(electrically EPROM,EEPROM)或闪存。存储器330也可以是易失性存储器,易失性存储器可以是随机存取存储器(random access memory,RAM),其用作外部高速缓存。通过示例性但不是限制性说明,许多形式的RAM可用,例如静态随机存取存储器(static RAM,SRAM)、动态随机存取存储器(dynamic RAM,DRAM)、同步动态随机存取存储器(synchronous DRAM,SDRAM)、双倍数据速率同步动态随机存取存储器(double data rate SDRAM,DDR SDRAM)、增强型同步动态随机存取存储器(enhanced SDRAM,ESDRAM)、同步连接动态随机存取存储器(synchlink DRAM,SLDRAM)和直接内存总线随机存取存储器(direct rambus RAM,DR RAM)。存储器330也可用于存储程序代码和数据,以便于处理器310调用存储器330中存储的程序代码执行上述方法实施例中的部分或者全部操作步骤,或者执行上述设备实施例中的相应功能。此外,计算设备300可能包含相比于图3展示的更多或者更少的组件,或者有不同的组件配置方式。

[0055] 总线340可以是快捷外围部件互连标准(peripheral component interconnect express,PCIe)总线,或扩展工业标准结构(extended industry standard architecture,EISA)总线、统一总线(unified bus,Ubus或UB)、计算机快速链接(compute express link,CXL)、缓存一致互连协议(cache coherent interconnect for accelerators,CCIX)等。总线340可以分为地址总线、数据总线、控制总线等。总线340除包括数据总线之外,还可以包括电源总线、控制总线和状态信号总线等。但是为了清楚说明起见,图3中仅用一条粗线表示,但并不表示仅有一根总线或一种类型的总线。

[0056] 本申请实施例提供的方法和设备是基于同一发明构思的,由于方法及设备解决问题的原理相似,因此方法与设备的实施例、实施方式、示例或实现方式可以相互参见,其中重复之处不再赘述。本申请实施例还提供一种系统,该系统包括多个计算设备,每个计算设备的结构可以参照上述所描述的计算设备的结构。该系统可实现的功能或者操作可以参照上述方法实施例中的具体实现步骤和/或上述装置实施例中所描述的具体功能,在此不再赘述。

[0057] 本申请实施例还提供一种计算机可读存储介质,所述计算机可读存储介质中存储有计算机指令,当所述计算机指令在计算机设备(如一个或者多个处理器)上运行时可以实现上述方法实施例中的方法步骤。所述计算机可读存储介质的处理器在执行上述方法步骤的具体实现可参照上述方法实施例中所描述的具体操作和/或上述装置实施例中所描述的具体功能,在此不再赘述。

[0058] 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。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。本申请实施例可以全部或部分地通过软件、硬件、固件或其他任意组合来实现。当使用软件实现时,上述实施例可以全部或部分地以计算机程序产品的形式实现。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上实施的计算机程序产品的形式。所述计算机程序产品包括一个或多个计算机指令。在计算机上加载或执行所述计算机程序指令时,全部或部分地产生按照本申请实施例所述的流程或功能。所述计算机可以为通用计算机、专用计算机、计算机网络、或者其他可编程装置。计算机指令可以存储在计算机可读存储介质中,或者从一个计算机可读存储介质向另一个计算机可读存储介质传输,例如,计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线(例如同轴电缆、光纤、数字用户线)或无线(例如红外、无线、微波等)方式向另一个网站站点、计算机、服务器或数据中心进行传输。计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集合的服务器、数据中心等数据存储设备。可用介质可以是磁性介质(如软盘、硬盘、磁带)、光介质、或者半导体介质。半导体介质可以是固态硬盘,也可以是随机存取存储器,闪存,只读存储器,可擦可编程只读存储器,电可擦可编程只读存储器,寄存器或任何其他形式的合适存储介质。

[0059] 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。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。

[0060] 在上述实施例中,对各个实施例的描述都各有侧重,某个实施例中未详述的部分,可以参见其它实施例的相关描述。显然,本领域的技术人员可以对本申请实施例进行各种改动和变型而不脱离本申请实施例的精神和范围。本申请实施例方法中的步骤可以根据实际需要进行顺序调整、合并或删减;本申请实施例系统中的模块可以根据实际需要进行划分、合并或删减。如果本申请实施例的这些修改和变型属于本申请权利要求及其等同技

术的范围之内,则本申请也意图包含这些改动和变型在内。

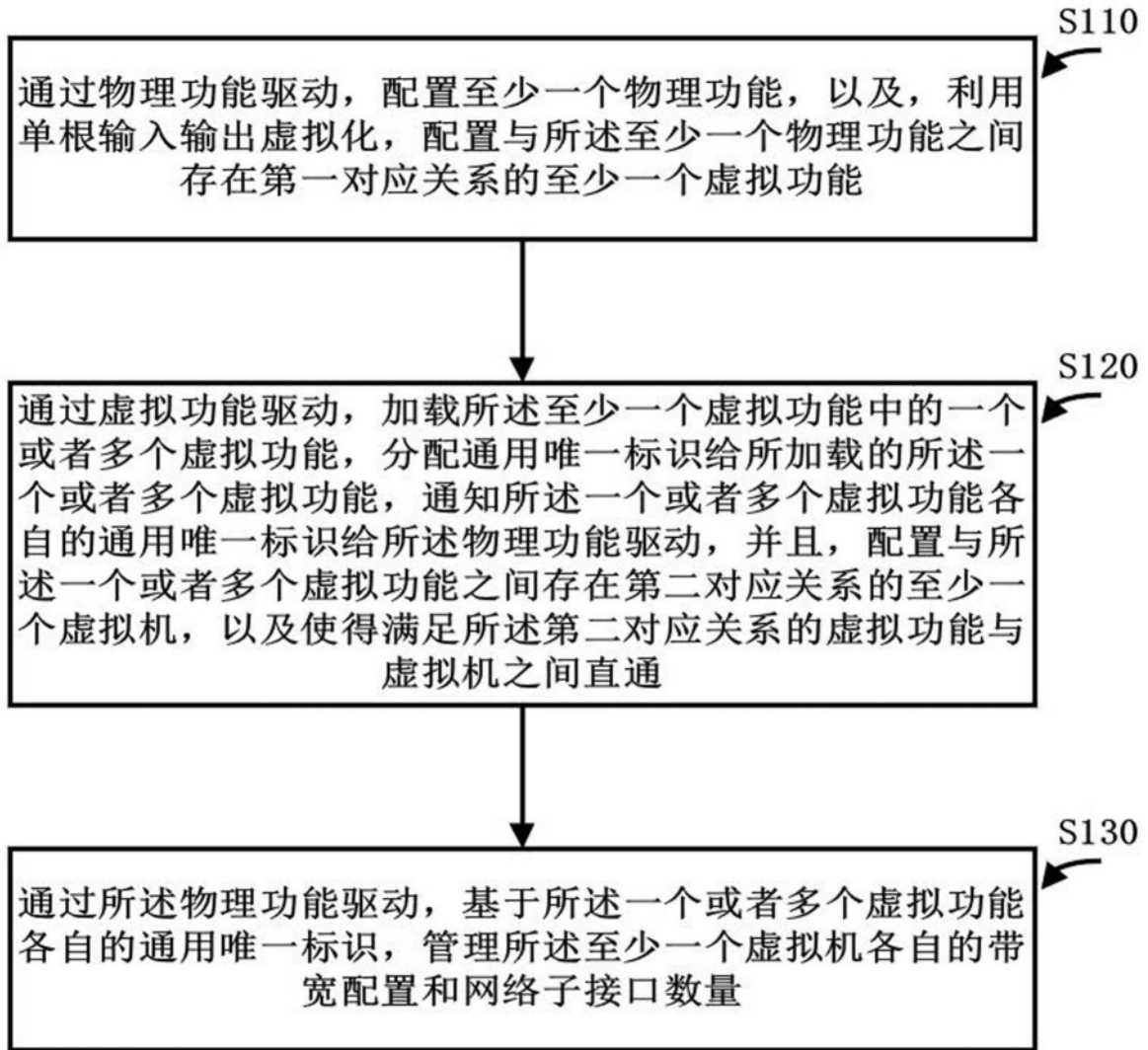


图1

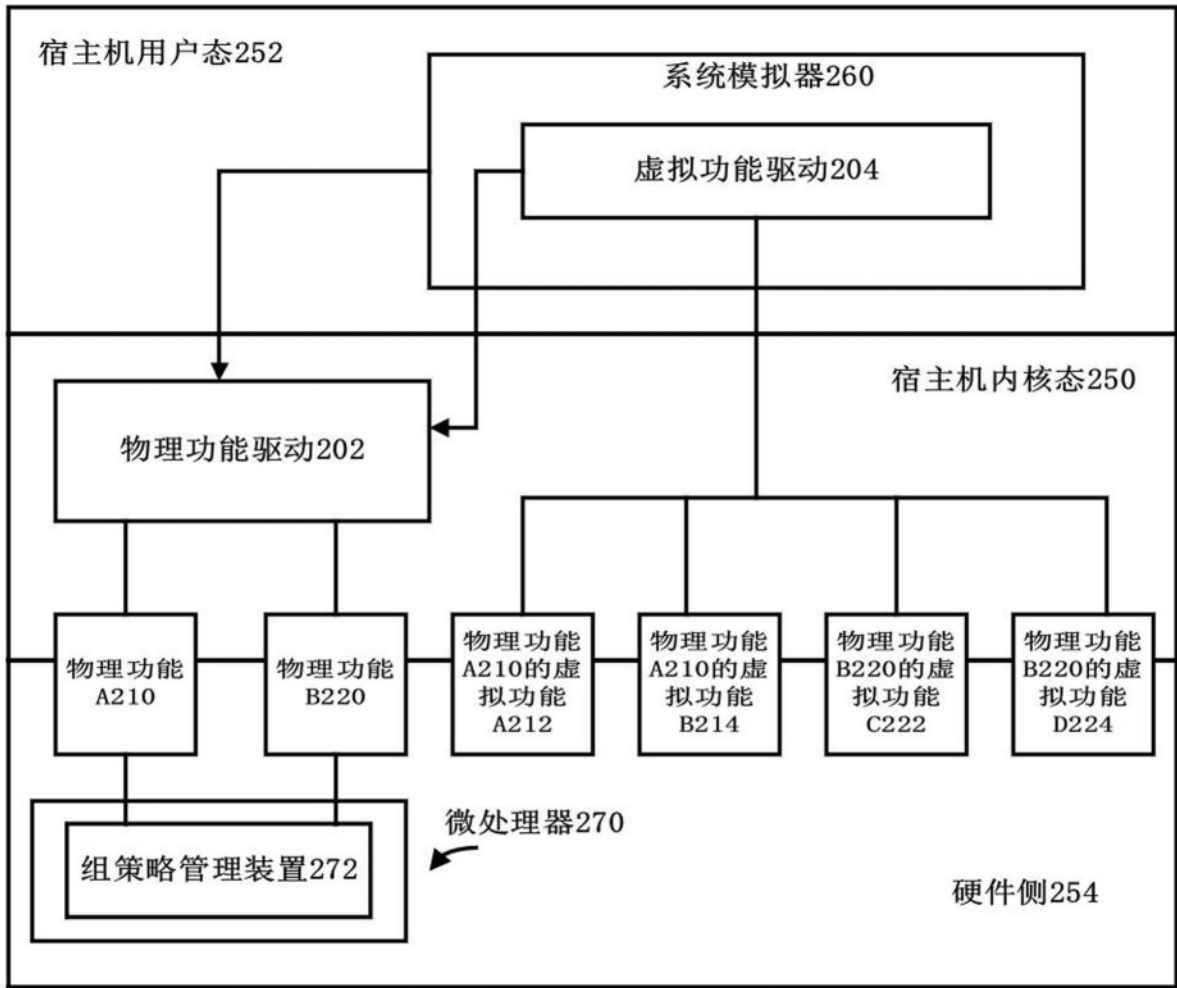


图2

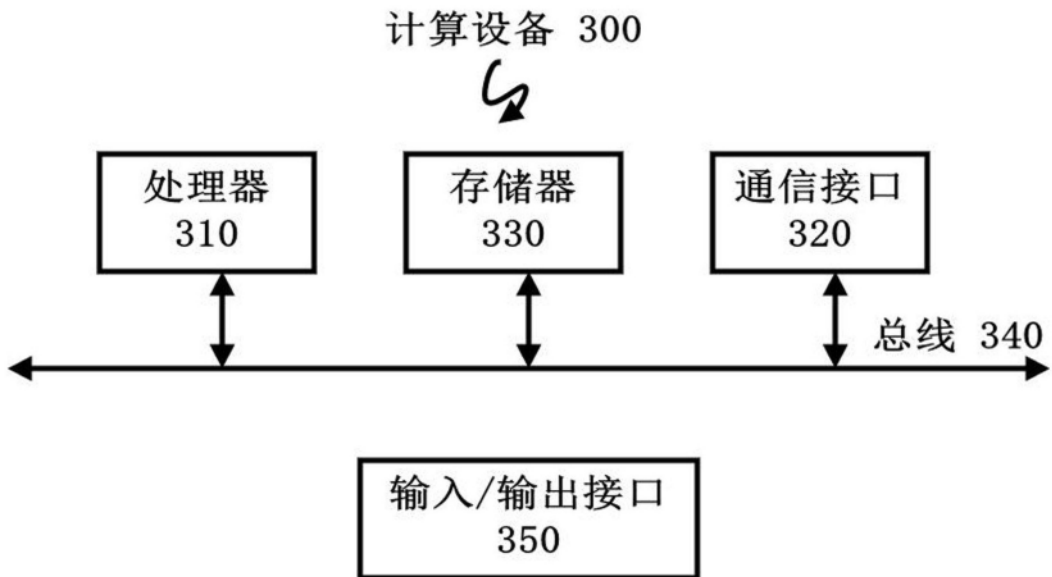


图3