

(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

(43) 国际公布日
2021年2月4日 (04.02.2021)



(10) 国际公布号
WO 2021/018267 A1

- (51) 国际专利分类号:
G06F 13/28 (2006.01)
- (21) 国际申请号: PCT/CN2020/106044
- (22) 国际申请日: 2020年7月31日 (31.07.2020)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910704642.2 2019年7月31日 (31.07.2019) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 曹水(CAO, Shui); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 刘耀崇(LIU, Yaochong); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 邱军(QIU, Jun); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 欧阳长春(OUYANG, Changchun); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (74) 代理人: 深圳市深佳知识产权代理事务所(普通合伙) (SHENPAT INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市罗湖区南湖街道春风路庐山大厦B座18C2、18D、18E、18E2, Guangdong 518001 (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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, SA, SC, SD, SE, SG, SK, SL,

(54) Title: LIVE MIGRATION METHOD FOR VIRTUAL MACHINE AND COMMUNICATION DEVICE

(54) 发明名称: 虚拟机热迁移方法和通信设备

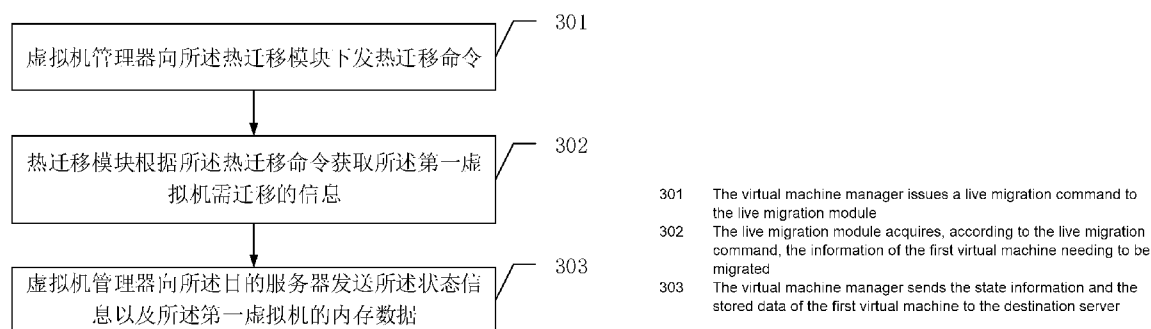


图 3

(57) Abstract: Disclosed in the embodiments of the present application is a live migration method for a virtual machine, capable of implementing live migration of the state of a single-root I/O virtualization (SR-IOV) network card. The method is applied to a source server in which a first virtual machine is running, and is used for migrating the first virtual machine to a destination server. The source server comprises an SR-IOV network card. The first virtual machine, a first virtual network card used by the first virtual machine, a virtual machine manager, and a live migration module are running in the source server. The first virtual network card obtained by virtualization based on the SR-IOV network card. The method comprises: the virtual machine manager issues a live migration command to the live migration module; the live migration module acquires, according to the live migration command, the information of the first virtual machine needing to be migrated, the information of the first virtual machine needing to be migrated comprising state information of the first virtual network card and stored data of the first virtual machine, and reports the information to the virtual machine manager; the virtual machine manager sends the state information and the stored data of the first virtual machine to the destination server.

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, WS, 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, ST, 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, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本申请实施例公开了一种虚拟机热迁移方法, 可实现SR-IOV网卡状态热迁移。应用于运行有第一虚拟机的源服务器, 该方法用于向目的服务器迁移该第一虚拟机, 该源服务器包括单根输入/输出虚拟化SR-IOV网卡, 该源服务器中运行有第一虚拟机、该第一虚拟机使用的第一虚拟网卡、虚拟机管理器和热迁移模块, 该第一虚拟网卡是基于该SR-IOV网卡虚拟出的, 该方法包括: 该虚拟机管理器向该热迁移模块下发热迁移命令; 该热迁移模块根据该热迁移命令获取该第一虚拟机需迁移的信息, 该第一虚拟机需迁移的信息包括该第一虚拟网卡的状态信息以及该第一虚拟机的内存数据, 并上报该虚拟机管理器; 该虚拟机管理器向该目的服务器发送该状态信息以及该第一虚拟机的内存数据。

虚拟机热迁移方法和通信设备

本申请要求于 2019 年 7 月 31 日提交中国专利局、申请号为 201910704642.2、发明名称为“虚拟机热迁移方法和通信设备”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

本申请涉及通信技术领域，特别涉及一种虚拟机热迁移方法和通信设备。

背景技术

虚拟机的热迁移又称动态迁移或在线迁移，是指在保证应用服务正常运行的同时，将一个虚拟机从一台物理服务器迁移到另一台物理服务器上，为了保证迁移过程中客户机服务的可用性，迁移过程只能有非常短暂的停机时间。依靠虚拟机在不同的主机之间迁移，集群能够有效地进行资源的分配和调度。

单根输入/输出虚拟化（single-root I/O virtualization, SR-IOV）是一种输入输出（input/output, I/O）虚拟化技术，该技术可以在一个物理快捷外围部件互连标准（peripheral component interconnection express, PCIE）设备上虚拟化出多个虚拟的 PCIE 设备，称之为虚拟功能（virtual function, VF）。SR-IOV 网卡可以虚拟化出多个虚拟网卡，虚拟机通过虚拟网卡执行业务，每个虚拟机可以绕开虚拟机管理器，直接与其对应的一个或多个虚拟网卡进行网络 I/O 的数据交互，虚拟网卡的状态信息保存在网卡的存储介质中，虚拟机管理器无法获取 SR-IOV 网卡的存储介质中的状态信息。

虚拟机热迁移时，需要将虚拟机对应虚拟网卡的状态信息由源服务器迁移至目的服务器，由于虚拟网卡的状态信息保存在物理网卡的存储介质中，虚拟机管理器不感知 SR-IOV 网卡的存储介质，无法获取该状态信息，因此无法实现虚拟机对应的虚拟网卡的状态信息迁移。

发明内容

本申请实施例提供了一种虚拟机热迁移方法，可以实现 SR-IOV 网卡状态热迁移，从而使得使用 SR-IOV 网卡的物理机同样可以执行虚拟机的热迁移过程。

本申请实施例第一方面提供了一种虚拟机热迁移方法，应用于运行有第一虚拟机的源服务器，该方法用于向目的服务器迁移该第一虚拟机，该源服务器包括单根输入/输出虚拟化 SR-IOV 网卡，该源服务器中运行有第一虚拟机、该第一虚拟机使用的第一虚拟网卡、虚拟机管理器和热迁移模块，该第一虚拟网卡是基于该 SR-IOV 网卡虚拟出的，该方法包括：该虚拟机管理器向该热迁移模块下发热迁移命令；该热迁移模块根据该热迁移命令获取该第一虚拟机需迁移的信息，该第一虚拟机需迁移的信息包括该第一虚拟网卡的状态信息以及该第一虚拟机的内存数据，并上报该虚拟机管理器；该虚拟机管理器向该目的服务器发送该状态信息以及该第一虚拟机的内存数据。

当源服务器中的第一虚拟机从源服务器迁移至目的服务器时，源服务器获取热迁移请求，源服务器与目的主机建立连接，并根据热迁移请求向目的服务器发送热迁移消息，该热迁移消息用于指示运行于该主机设备的第一虚拟机迁移至该目的服务器；热迁移模块接收该虚拟

机管理器根据该热迁移请求下发的热迁移命令，可以通过新增的该热迁移模块获取该第一虚拟机对应的第一虚拟网卡的状态信息；然后，该虚拟机管理器将该状态信息发送给该目的服务器。由此，可以将该状态信息迁移至目的服务器，进而实现第一虚拟机的迁移。

在第一方面的一种可能的实现方式中，该热迁移模块根据该热迁移命令获取该第一虚拟机需迁移的信息包括：该热迁移模块根据该热迁移命令，将该状态信息拷贝至该虚拟机管理器分配的该第一内存空间，该第一内存空间不包括在给该第一虚拟机分配的虚拟内存内。

本申请实施例提供的虚拟机热迁移方法，热迁移模块可以将获取的状态信息拷贝至虚拟机管理器分配的第一内存空间，该第一内存空间不包括在给该第一虚拟机分配的虚拟内存内，第一虚拟机不感知该迁移过程。

在第一方面的一种可能的实现方式中，该热迁移模块根据该热迁移命令，将该状态信息拷贝至该虚拟机管理器分配的该第一内存空间包括：该热迁移模块从该 SR-IOV 网卡的基地址寄存器、微处理器和微码中拷贝该状态信息至该第一内存空间。

本申请实施例提供的虚拟机热迁移方法，热迁移模块获取该状态信息的包括了分别从物理网卡的基地址寄存器、微处理器和微码中进行拷贝，提供了获取状态信息的一种可行的实现方式。

在第一方面的一种可能的实现方式中，该热迁移模块根据该热迁移命令获取该第一虚拟机需迁移的信息之前，该方法还包括：该热迁移模块锁定该第一虚拟网卡的第一任务队列，以使得该第一任务队列无法接收新的任务，该第一任务队列用于放置为该第一虚拟机下发的任务；该第一虚拟机的第一虚拟功能驱动清空该第一任务队列中所有已有的任务。

在第一虚拟机的迁移过程中，若用户通过该第一虚拟机下发用户命令改变对应的虚拟网卡的状态信息，将可能发生在命令执行完成之前，第一虚拟机对应的虚拟网卡状态信息已经迁移到目标虚拟机的情况，有可能产生命令超时等异常，导致业务中断。本申请实施例提供的虚拟机热迁移方法，在获取虚拟网卡的状态信息之前，该热迁移模块可以锁定第一任务队列，第一虚拟功能驱动可执行第一任务队列中的任务直至排空该队列。这样用户将无法修改第一虚拟机对应的虚拟网卡的状态信息，还可以避免业务异常。

在第一方面的一种可能的实现方式中，该虚拟机管理器向该热迁移模块下发热迁移命令之后，该方法还包括：该热迁移模块对该内存数据进行迭代拷贝；在剩余的该内存数据可一次拷贝完成的情况下，该第一虚拟功能驱动清空该第一任务队列中所有已有的任务之后，该虚拟机管理器挂起该第一虚拟机；该热迁移模块执行最后一次内存数据拷贝。

本申请实施例提供的虚拟机热迁移方法，限定了热迁移模块在内存数据迭代拷贝剩余内存可一次拷贝完成时，锁定并排空执行用户命令的通道，由于锁定执行用户命令的通道后，将不再响应新的用户命令，剩余内存可一次拷贝完成时执行锁定操作，可以减少不响应用户命令阶段的时长。

本申请实施例第二方面提供了一种虚拟机热迁移方法，应用于虚拟机迁移的目的服务器，该目的服务器包括 SR-IOV 网卡，该目的服务器中运行有虚拟机管理器和热迁移模块，该方法包括：该虚拟机管理器获取热迁移命令，该热迁移命令用于指示源服务器的第一虚拟机将迁移至该目的服务器的第二虚拟机；该虚拟机管理器接收该源服务器发送的该第一虚拟机对应的第一虚拟网卡的状态信息；该热迁移模块根据该虚拟机管理器的指示，将该状态信息恢

复至该 SR-IOV 网卡。

当源服务器中的第一虚拟机从源服务器迁移至目的服务器的第二虚拟机时，需要将第一虚拟机对应的第一网卡的状态信息也迁移至目的服务器的物理网卡中。该虚拟机管理器接收源服务器发送的第一虚拟机对应的虚拟网卡的状态信息，目的服务器的虚拟机管理器接收到
5 该状态信息后，通过该热迁移模块需要将该状态信息恢复至物理网卡中。使得第一虚拟机的成功迁移。

在第二方面的一种可能的实现方式中，该热迁移模块根据该虚拟机管理器的指示，将该状态信息恢复至该 SR-IOV 网卡包括：该热迁移模块将该状态信息分别恢复至该 SR-IOV 网卡的基地址寄存器、微码和微处理器中。

10 本申请实施例提供的虚拟机热迁移方法，可以将状态信息分别写入该物理网卡的基地址寄存器、微码和微处理器中，增强了方案的可实现性。

在第二方面的一种可能的实现方式中，该热迁移模块将该状态信息分别恢复至该 SR-IOV 网卡的基地址寄存器、微码和微处理器中包括：该热迁移模块将该状态信息中的第一信息拷贝至该微处理器，该第一信息为从该源服务器的只读基地址寄存器中获取的信息，该微处理
15 器用于将该第一信息恢复至该 SR-IOV 网卡的只读基地址寄存器。

对于从源服务器的只读基地址寄存器中获取的第一信息，热迁移模块不能直接写入目的服务器的只读基地址寄存器，而是通过拷贝至微处理器中进行恢复，由于微处理器权限较高，可以解决只读基地址寄存器中不能直接写入的问题。

在第二方面的一种可能的实现方式中，该虚拟机管理器接收该源服务器发送的该第一虚拟机对应的第一虚拟网卡的状态信息之前，该方法还包括：该虚拟机管理器通过该热迁移模块分配该第二虚拟机对应的第二虚拟网卡的标识信息。

本申请实施例提供的虚拟机热迁移方法，在接收源服务器发送的第一虚拟机对应的虚拟网卡的状态信息之前，进行全局资源分配，可以避免全局资源被占用带来的冲突。

在第二方面的一种可能的实现方式中，该虚拟机管理器接收该源服务器发送的该第一虚拟机对应的第一虚拟网卡的状态信息包括：该虚拟机管理器接收该源服务器发送的第二信息，该第二信息为该第二虚拟机对应的第二虚拟网卡的第二任务队列的锁定状态信息，该第二虚拟网卡是基于该 SR-IOV 网卡虚拟出的，该第二任务队列用于放置为该第二虚拟机下发的任务；
25 该方法还包括：该热迁移模块解锁该第二任务队列。

本申请实施例提供的虚拟机热迁移方法，可以将锁定的执行用户命令的通道解锁，这样
30 可以在网卡状态恢复后继续执行用户下发的命令。

该方法还包括：该虚拟机管理器接收该源服务器发送的第一虚拟机对应的内存数据；该虚拟机管理器恢复该第一虚拟机对应的内存数据。

本申请实施例第三方面提供了一种虚拟机热迁移方法，应用于运行有第一虚拟机的源服务器，该方法用于向目的服务器迁移该第一虚拟机，该源服务器包括单根输入/输出虚拟化
35 SR-IOV 网卡，该源服务器中运行有第一虚拟机、该第一虚拟机使用的第一虚拟网卡、虚拟机管理器和物理功能驱动，该第一虚拟网卡是基于该 SR-IOV 网卡虚拟出的，该方法包括：该虚拟机管理器向该虚拟功能驱动下发热迁移命令；该虚拟功能驱动根据该热迁移命令获取该第一虚拟机需迁移的信息，该第一虚拟机需迁移的信息包括该第一虚拟机对应的第一虚拟网卡

的状态信息以及该第一虚拟机的内存数据，并上报该虚拟机管理器；该虚拟机管理器向该目的服务器发送该状态信息以及该第一虚拟机的内存数据。

本申请实施例提供的虚拟机热迁移方法，虚拟机管理器不感知 SR-IOV 网卡的存储介质，无法获取虚拟网卡的状态信息，物理功能驱动则可以与 SR-IOV 网卡进行交互，本申请实施例中虚拟机管理器通过 VFIO 与物理功能驱动连接，进而可以通过物理功能驱动获取虚拟网卡的状态信息，进而实现第一虚拟机的迁移。

在第三方面的一种可能的实现方式中，该虚拟功能驱动根据该热迁移命令获取该第一虚拟机需迁移的信息包括：该虚拟功能驱动根据该热迁移命令，将该状态信息拷贝至该虚拟机管理器分配的该第一内存空间，该第一内存空间不包括在给该第一虚拟机分配的虚拟内存内。

在第三方面的一种可能的实现方式中，该虚拟功能驱动根据该热迁移命令，将该状态信息拷贝至该虚拟机管理器分配的该第一内存空间包括：该虚拟功能驱动从该 SR-IOV 网卡的基地址寄存器、微处理器和微码中拷贝该状态信息至该第一内存空间。

在第三方面的一种可能的实现方式中，该虚拟功能驱动根据该热迁移命令，将该状态信息拷贝至该虚拟机管理器分配的该第一内存空间之前，该方法还包括：该虚拟功能驱动锁定该第一虚拟网卡的第一任务队列，以使得该第一任务队列无法接收新的任务，该第一任务队列用于放置为该第一虚拟机下发的任务；该第一虚拟机运行的第一虚拟功能驱动清空该第一任务队列中所有已有的任务。

在第三方面的一种可能的实现方式中，该虚拟机管理器获取热迁移请求之后，该方法还包括：该虚拟功能驱动对该内存数据进行迭代拷贝；在剩余的该内存数据可一次拷贝完成的情况下，该第一虚拟功能驱动清空该第一任务队列中所有已有的任务之后，该虚拟机管理器挂起该第一虚拟机；该虚拟功能驱动执行最后一次内存数据拷贝。

在第三方面的一种可能的实现方式中，该虚拟功能驱动通过用户态驱动框架 VFIO 与虚拟机管理器连接。

在第三方面的一种可能的实现方式中，该虚拟功能驱动包括热迁移模块，该热迁移模块用于根据虚拟机管理器的指示获取第一虚拟网卡的状态信息。

本申请实施例提供的虚拟机热迁移方法，在物理功能驱动可以通过新增的热迁移模块，用于辅助实现虚拟机热迁移。虚拟机管理器不感知 SR-IOV 网卡的存储介质，无法获取虚拟网卡的状态信息，物理功能驱动则可以与 SR-IOV 网卡进行交互，本申请实施例中通过在物理功能驱动中新增热迁移模块，虚拟机管理器通过 VFIO 与该热迁移模块连接，进而可以通过热迁移模块获取虚拟网卡的状态信息，进而实现第一虚拟机的迁移。

本申请实施例第四方面提供了一种虚拟机热迁移方法，其特征在于，应用于虚拟机迁移的目的服务器，该目的服务器包括 SR-IOV 网卡，该目的服务器中运行有虚拟机管理器和物理功能驱动，该方法包括：该虚拟机管理器获取热迁移命令，该热迁移命令用于指示源服务器的第一虚拟机将迁移至该目的服务器的第二虚拟机；该虚拟机管理器接收该源服务器发送的该第一虚拟机对应的第一虚拟网卡的状态信息；该虚拟功能驱动根据该虚拟机管理器的指示，将该状态信息恢复至该 SR-IOV 网卡。

在第四方面的一种可能的实现方式中，该虚拟功能驱动根据该虚拟机管理器的指示，将该状态信息恢复至该 SR-IOV 网卡包括：该虚拟功能驱动将该状态信息分别恢复至该 SR-IOV

网卡的基地址寄存器、微码和微处理器中。

在第四方面的一种可能的实现方式中，该虚拟功能驱动将该状态信息分别恢复至该 SR-IOV 网卡的基地址寄存器、微码和微处理器中包括：该虚拟功能驱动将该状态信息中的第一信息拷贝至该微处理器，该第一信息为从该源服务器的只读基地址寄存器中获取的信息；该
5 微处理器将该第一信息恢复至该 SR-IOV 网卡的只读基地址寄存器。

在第四方面的一种可能的实现方式中，该虚拟机管理器接收该源服务器发送的该第一虚拟机对应的第一虚拟网卡的状态信息之前，该方法还包括：该虚拟机管理器通过该虚拟功能驱动分配该第二虚拟机对应的第二虚拟网卡的标识信息。

该虚拟机管理器接收该源服务器发送的该第一虚拟机对应的第一虚拟网卡的状态信息包
10 括：该虚拟机管理器接收该源服务器发送的第二信息，该第二信息为该第二虚拟机对应的第二虚拟网卡的第二任务队列的锁定状态信息，该第二虚拟网卡是基于该 SR-IOV 网卡虚拟出的，该第二任务队列用于放置为该第二虚拟机下发的任务；该方法还包括：该虚拟功能驱动解锁该第二任务队列。

在第四方面的一种可能的实现方式中，该虚拟功能驱动包括热迁移模块，该热迁移模块
15 用于根据虚拟机管理器的指示将该状态信息恢复至该 SR-IOV 网卡。

本申请实施例提供的虚拟机热迁移方法，在物理功能驱动可以通过新增的热迁移模块用于辅助实现虚拟机热迁移。

本申请实施例第五方面提供了一种通信设备，该通信设备为虚拟机热迁移的源源端，该通信设备中运行有第一虚拟机、该第一虚拟机使用的第一虚拟网卡、虚拟机管理器和热迁移
20 模块，该第一虚拟网卡是基于 SR-IOV 网卡虚拟出的，该方法包括：该虚拟机管理器，用于向该热迁移模块下发热迁移命令；该热迁移模块，用于根据该热迁移命令获取该第一虚拟机需迁移的信息，该第一虚拟机需迁移的信息包括该第一虚拟网卡的状态信息以及该第一虚拟机的内存数据，并上报该虚拟机管理器；该虚拟机管理器，还用于向该目的端通信设备发送该状态信息以及该第一虚拟机的内存数据。

在第五方面的一种可能的实现方式中，该热迁移模块具体用于：根据该热迁移命令，将该状态信息拷贝至该虚拟机管理器分配的该第一内存空间，该第一内存空间不包括在给该第一虚拟机分配的虚拟内存内。

在第五方面的一种可能的实现方式中，该热迁移模块具体用于：从该 SR-IOV 网卡的基地址寄存器、微处理器和微码中拷贝该状态信息至该第一内存空间。

在第五方面的一种可能的实现方式中，该热迁移模块，还用于锁定该第一虚拟网卡的第一任务队列，以使得该第一任务队列无法接收新的任务，该第一任务队列用于放置为该第一虚拟机下发的任务；该通信设备还包括：该第一虚拟机的第一虚拟功能驱动，用于清空该第一任务队列中所有已有的任务。

在第五方面的一种可能的实现方式中，该热迁移模块还用于：对该内存数据进行迭代拷贝；在剩余的该内存数据可一次拷贝完成的情况下，该第一虚拟功能驱动清空该第一任务队列中所有已有的任务之后，该虚拟机管理器还用于，挂起该第一虚拟机；该热迁移模块，还用于执行最后一次内存数据拷贝。

本申请实施例第六方面提供了一种通信设备，该通信设备为虚拟机热迁移的目的端通信

设备，该通信设备包括 SR-IOV 网卡，该通信设备中运行有虚拟机管理器和热迁移模块，该通信设备包括：该虚拟机管理器，用于获取热迁移命令，该热迁移命令用于指示源端通信设备的第一虚拟机将迁移至该目的端通信设备的第二虚拟机；该虚拟机管理器，还用于接收该源端通信设备发送的该第一虚拟机对应的第一虚拟网卡的状态信息；该热迁移模块，还用于根据该虚拟机管理器的指示，将该状态信息恢复至该 SR-IOV 网卡。

在第六方面的一种可能的实现方式中，该热迁移模块具体用于：将该状态信息分别恢复至该 SR-IOV 网卡的基地址寄存器、微码和微处理器中。

在第六方面的一种可能的实现方式中，该热迁移模块具体用于：将该状态信息中的第一信息拷贝至该微处理器，该第一信息为从该源通信设备的只读基地址寄存器中获取的信息，该微处理器用于将该第一信息恢复至该 SR-IOV 网卡的只读基地址寄存器。

在第六方面的一种可能的实现方式中，该虚拟机管理器还用于：通过该热迁移模块分配该第二虚拟机对应的第二虚拟网卡的标识信息。

在第六方面的一种可能的实现方式中，该虚拟机管理器具体用于：接收该源通信设备发送的第二信息，该第二信息为该第二虚拟机对应的第二虚拟网卡的第二任务队列的锁定状态信息，该第二虚拟网卡是基于该 SR-IOV 网卡虚拟出的，该第二任务队列用于放置为该第二虚拟机下发的任务；该热迁移模块，还用于解锁该第二任务队列。

本申请实施例第七方面提供了一种通信设备，为虚拟机热迁移的源端通信设备，包括：处理器、存储器和 SR-IOV 网卡；该处理器用于运行第一虚拟机、虚拟机管理器和热迁移模块，该第一虚拟机运行第一虚拟功能驱动，该处理器通过该虚拟机管理器、该热迁移模块和该第一虚拟功能驱动执行如前述第一方面提供的各实施方式的方法。

本申请实施例第七方面提供的一种通信设备中，该处理器、存储器和 SR-IOV 网卡电连接，可选地，该通信设备还包括总线，该处理器、存储器和 SR-IOV 网卡通过该总线相连。

本申请实施例第八方面提供了一种通信设备，为虚拟机热迁移的目的端通信设备，包括：处理器、存储器和 SR-IOV 网卡；该处理器用于运行虚拟机管理器和热迁移模块，并通过该虚拟机管理器和热迁移模块执行如前述第二方面提供的各实施方式的方法。

本申请实施例第八方面提供的一种通信设备中，该处理器、存储器和 SR-IOV 网卡电连接，可选地，该通信设备还包括总线，该处理器、存储器和 SR-IOV 网卡通过该总线相连。

本申请实施例第九方面提供了一种通信设备，为虚拟机热迁移的源端通信设备，包括：处理器、存储器和 SR-IOV 网卡；该处理器用于运行第一虚拟机、虚拟机管理器和物理功能驱动，该第一虚拟机运行第一虚拟功能驱动，该处理器通过该虚拟机管理器、该物理功能驱动和该第一虚拟功能驱动执行前述第三方面提供的各实施方式的方法。

本申请实施例第九方面提供的一种通信设备中，该处理器、存储器和 SR-IOV 网卡电连接，可选地，该通信设备还包括总线，该处理器、存储器和 SR-IOV 网卡通过该总线相连。

本申请实施例第十方面提供了一种通信设备，为虚拟机热迁移的目的端通信设备，包括：处理器、存储器和 SR-IOV 网卡；该处理器用于运行虚拟机管理器，和物理功能驱动，并通过该虚拟机管理器和该物理功能驱动执行前述第四方面提供的各实施方式的方法。

本申请实施例第十方面提供的一种通信设备中，该处理器、存储器和 SR-IOV 网卡电连接，可选地，该通信设备还包括总线，该处理器、存储器和 SR-IOV 网卡通过该总线相连。

本申请实施例第十一方面提供了一种包含指令的计算机程序产品，当其在计算机上运行时，使得该计算机执行前述第一方面至第四方面提供的各实施方式的步骤。

本申请实施例第十二方面提供了一种计算机可读存储介质，包括指令，当该指令在计算机上运行时，使得计算机执行前述第一方面至第四方面提供的各实施方式的步骤。

5 本申请实施例第十三方面提供了一种虚拟机热迁移系统，包括前述第五方面提供的源端通信设备，和前述第六方面提供的目的端通信设备。

从以上技术方案可以看出，本申请实施例具有以下优点：

本申请实施例提供的虚拟机热迁移方法，通过热迁移模块可获取 SR-IOV 网卡中的第一虚拟网卡的状态信息，并将该状态信息迁移至目的服务器中。本申请实施例提供的虚拟机热迁移方法可以使得该状态信息迁移到第二虚拟机侧，进而保证第一虚拟机的成功迁移。

10

附图说明

图 1 为虚拟机热迁移的一种迁移架构示意图；

图 2 为本申请实施例的虚拟机热迁移的系统架构示意图；

15 图 3 为本申请实施例中虚拟机热迁移方法的一个实施例示意图；

图 4 为本申请实施例中虚拟机热迁移方法的另一个实施例示意图；

图 5 为本申请实施例中虚拟机热迁移方法的一个实施例交互示意图；

图 6 为本申请实施例中锁定并排空任务队列的一个实施例示意图；

图 7 为本申请实施例中源端通信设备的一个实施例示意图；

20 图 8 为本申请实施例中目的端通信设备的一个实施例示意图；

图 9 为本申请实施例中源端通信设备的另一个实施例示意图；

图 10 为本申请实施例中目的端通信设备的另一个实施例示意图。

具体实施方式

本申请实施例提供了一种虚拟机热迁移方法，可以实现网卡的状态信息的迁移，从而使
25 得使用 SR-IOV 网卡的物理机同样可以执行虚拟机的热迁移过程。

请参阅图 1，为虚拟机热迁移的一种迁移架构示意图。

该迁移架构中包括虚拟机迁移前的源服务器 100 和虚拟机迁移后的目的服务器 110。在一种可能的应用场景中，可以将运行在源服务器 100 中的虚拟机 1 的运行状态迁移到目的服务器 110 中任一虚拟机，例如，将虚拟机 1 迁移至虚拟机 2。

30 源服务器和目的服务器均为部署了虚拟机管理器的宿主机，可以为虚拟机提供物理资源。通常在一个宿主机上可以创建多个虚拟机，创建虚拟机的个数取决于宿主机本身的硬件配置和需要创建的虚拟机规格大小。此处对于源服务器 100 上运行的虚拟机数量 N 和目的服务器上运行的虚拟机数量 M 不做限定。在实际应用场景中，宿主机可以是提供各种云服务的服务器。

35 请参阅图 2，为本申请实施例的虚拟机热迁移的系统示意图。

该系统中包括虚拟机迁移前的源服务器和虚拟机热迁移后的目的服务器。

源服务器包括主机和 SR-IOV 网卡，具体地，主机 (HOST) 中运行至少一个虚拟机，图 2 中仅示出一个热迁移前的第一虚拟机 VM1。

该主机还运行有源服务器的虚拟机管理器，该虚拟机管理器可以管理一个或多个虚拟机 VM，每个虚拟机 VM 上可以运行操作系统以及运行在操作系统上的各种应用。

该虚拟机管理器为可以是运行在主机设备上的管理模块，例如 hypervisor 或虚拟机监控器 (virtual machine monitor, VMM)，也可以由定制的逻辑芯片实现。该虚拟机管理器负责虚拟机的管理和控制，包括虚拟机的创建、迁移、上电和下电等操作。该虚拟机管理器具体可以是 qemu (Quick EMUlator)、VMWare、ESXi 或基于内核的虚拟机(kernel-based virtual machine, KVM)等，具体此处不做限定。

热迁移模块，是本实施例中新增模块，为虚拟机管理器提供热迁移接口，可以根据虚拟机管理器的指示辅助实现热迁移，本申请实施例中热迁移模块可以用于获取虚拟机对应的虚拟网卡的状态信息。可选的，在物理功能驱动中新增该热迁移模块，由于虚拟机管理器不感知 SR-IOV 网卡的存储介质，无法获取虚拟网卡的状态信息，物理功能驱动则可以与 SR-IOV 网卡进行交互。虚拟机管理器通过 VFIO 与该热迁移模块连接，可以通过热迁移模块获取虚拟网卡的状态信息，进而实现第一虚拟机的迁移。

SR-IOV 网卡包括至少一个物理功能(physical function, PF)以及至少一个虚拟功能 VF，本实施例中 SR-IOV 网卡的虚拟功能即虚拟网卡，图 1 中仅示出了一个物理功能与一个第一虚拟网卡。一个虚拟机对应至少一个虚拟网卡，虚拟网卡的状态信息可以保存在 SR-IOV 网卡的基地址寄存器 (base address register, BAR)、管理单元和数据处理单元中。本实施例中管理单元包括微处理器，数据处理单元包括微码。

该主机设备上还运行有管理 SR-IOV 网卡的 PF 驱动，该 PF 驱动可以直接访问 PF 的所有资源，并负责配置和管理所有 VF，每个虚拟机 VM 运行有 VF 驱动来管理各自对应的 VF。

目的服务器包括主机 (HOST) 和 SR-IOV 网卡，该 HOST 上运行有目的服务器的虚拟机管理器和热迁移模块，主机设备与 SR-IOV 网卡连接。

目的服务器的热迁移模块，为本实施例中新增模块，为虚拟机管理器提供热迁移接口，可以根据虚拟机管理器的指示辅助实现热迁移，本申请实施例中热迁移模块可以用于将获取的虚拟机对应的虚拟网卡的状态信息恢复至 SR-IOV 网卡。可选的，在物理功能驱动中新增该热迁移模块，虚拟机管理器通过 VFIO 与该热迁移模块连接。

可以理解的是，在实际应用中，一个服务器既可能作为虚拟机迁移的源端，也可以作为虚拟机迁移的目的端，因此本实施例中提供的源服务器和目的服务器的软件模块的功能可以集成在一个服务器中实现。

需要说明的是，尽管图 2 中未示出，但虚拟机管理器和各虚拟机之间存在有通信链路。尽管图 2 中未示出，该主机设备还包括与运行虚拟机 VM 相关的硬件，例如中央处理器 (central processing unit, CPU) 及至少一个计算机可读存储介质的存储器，包括硬盘或内存等。除此之外，该主机设备还可以包括通信单元、输入单元、显示单元等。

由于每个虚拟机可以与其对应的一个或多个虚拟网卡直接进行网络 I/O 的数据交互，无需通过虚拟机管理器，既降低了虚拟机管理器中的 I/O 性能开销，又保证了每个 VM 的 I/O 吞吐效率。具体的，虚拟机 VM 从虚拟网卡收发数据报文时，可实现收发完全独立，在进行直接内存访问 (direct memory access, DMA) 写入的过程中虚拟机管理器并不能感知。虚拟机对应的虚拟网卡的状态信息存储在 SR-IOV 网卡的存储介质中，虚拟机管理器也不能感知。

虚拟机热迁移中的关键技术点包括内存状态迁移和网卡状态迁移。

通常虚拟机内存的迁移可为三个阶段：

迭代预拷贝阶段：虚拟机迁移过程启动后，其依然运行在源服务器上，虚拟机业务不中断，首先把虚拟机的全部内存复制到目的服务器上，然后向目的服务器不停的迭代复制虚拟机中变化的内存数据。当迭代到需要复制的内存小到一定的门限后，进入停机拷贝阶段。

停机拷贝阶段：虚拟机停机，业务中断，拷贝虚拟机的残余内存数据。

恢复阶段：目的虚拟机完成运行前的恢复处理，启动目的虚拟机，整个迁移过程完成。

虚拟机内存的迁移方式有多种，本申请实施例中并不限定。

网卡状态迁移是指将进行迁移的虚拟机对应的虚拟网卡的状态信息由源服务器迁移至目的服务器。

虚拟机迁移过程中需进行迁移的虚拟网卡的状态信息主要包括以下几类：

(1)、设备信息，包括：状态信息和中断表等，存储在 BAR 寄存器的 BAR0/1 和 BAR2/3 空间中。

(2)、芯片与主机交互制信息和报文上下文，包括：响应事件队列、Mailbox、控制队列、报文描述队列和队列状态等。部分存储在 BAR 寄存器 BAR0/1 空间；部分存储在微码中网卡芯片内部表项中，可以通过微码获取和保存。

(3)、网卡业务上下文信息，主要用于管理 VF、PF 的功能，包括：函数表，MAC-VLAN 表和 RSS 模板表等，可以通过微处理器获取和保存。

由于虚拟网卡的状态信息存储在 SR-IOV 网卡的 BAR 寄存器、微码和微处理器中，当源服务器的第一虚拟机要迁移到目的服务器的第二虚拟机时，源服务器的虚拟机管理器无法获取 SR-IOV 网卡的存储介质中该第一虚拟机对应的虚拟网卡的状态信息。要实现第一虚拟机的成功迁移，该第一虚拟网卡的状态信息也需要迁移到第二虚拟机侧，以使得第二虚拟机对应的第二虚拟网卡恢复工作。为解决此问题，本申请实施例中通过新增的热迁移模块，可以获取网卡中的第一虚拟网卡的状态信息，虚拟机管理器可以将该状态信息拷贝至目的服务器，最终保证第一虚拟机的成功迁移。

基于图 2 提供的虚拟机热迁移的系统架构，请参阅图 3，为本申请实施例中虚拟机热迁移方法的一个实施例示意图。

301、虚拟机管理器向所述热迁移模块下发热迁移命令；

当源服务器中的第一虚拟机从源服务器迁移至目的服务器时，源服务器获取热迁移请求，源服务器根据该热迁移请求与目的主机建立连接，并求向该目的服务器发送热迁移请求，指示源服务器中第一虚拟机迁移至该目的服务器，目的服务器可以进行热迁移前的准备工作，例如，目的服务器为该第一虚拟机预保留虚拟机资源。

虚拟机管理器获取热迁移请求的方式有多种，可选的，用户通过源服务器发送热迁移请求，或者服务器根据预设的规则确定第一虚拟机进行热迁移，具体此处不做限定。

虚拟机管理器向所述热迁移模块下发热迁移命令，需要说明的是，虚拟机管理器在第一虚拟机迁移过程中可以向热迁移模块发送多个热迁移命令，此处对于热迁移命令的数量不做限定。可选的，虚拟机管理器通过用户态驱动框架（virtual function I/O，VFIO），在热迁移的不同阶段，向热迁移模块发送不同的热迁移命令。

302、热迁移模块根据所述热迁移命令获取所述第一虚拟机需迁移的信息；

第一虚拟机需迁移的信息包括该第一虚拟网卡的状态信息以及该第一虚拟机的内存数据，并上报所述虚拟机管理器。

其中，第一虚拟网卡的状态信息存储在 SR-IOV 网卡中，具体地，存储在 SR-IOV 网卡的 BAR 寄存器、管理单元和数据处理单元中。

虚拟机管理器根据该热迁移请求，通过该热迁移模块获取存储在 SR-IOV 网卡的 BAR 寄存器、管理单元和数据处理单元中的第一虚拟网卡的状态信息。

可选的，由虚拟机管理器分配的第一内存空间，热迁移模块根据状态信息存储位置分别从网卡的 BAR 寄存器、微处理器和微码中获取第一虚拟网卡的状态信息，并将该状态信息保存至该第一内存空间。需要说明的是，该第一内存空间不属于第一虚拟机对应的内存空间，由虚拟机管理器直接分配控制。

可选的，虚拟机管理器通过 VFIO 向热迁移模块发送控制命令。

303、虚拟机管理器向所述目的服务器发送所述状态信息以及所述第一虚拟机的内存数据；虚拟机管理器将步骤 302 中获取的第一虚拟机的内存数据发送给该目的服务器进行恢复，此外，虚拟机管理器还将步骤 302 中获取的该状态信息发送给该目的服务器，以使得目的服务器可以根据该状态信息对迁移后的第二虚拟机进行网卡状态恢复，进而实现热迁移。

本申请实施例中通过源服务器通过新增的热迁移模块，可以获取网卡中的第一虚拟网卡的状态信息，并将该状态信息迁移至第二虚拟机。使得这些状态信息能够迁移到第二虚拟机侧，最终保证第一虚拟机的成功迁移。

基于图 2 提供的虚拟机热迁移的系统架构，请参阅图 4，为本申请实施例中虚拟机热迁移方法的另一个实施例示意图。

401、虚拟机管理器接收源服务器发送的第一虚拟机对应的虚拟网卡的状态信息；

当源服务器中的第一虚拟机从源服务器迁移至目的服务器的第二虚拟机时，需要将第一虚拟机对应的第一网卡的状态信息也迁移至目的服务器的 SR-IOV 网卡中。

虚拟机管理器接收源服务器发送的该状态信息。

402、虚拟机管理器通过该热迁移模块，将该状态信息恢复至该 SR-IOV 网卡；

目的服务器的虚拟机管理器接收到该状态信息后，需要将该状态信息分别恢复至 SR-IOV 网卡的 BAR 寄存器、微码和微处理器中。

从源服务器的微处理器和微码获取的该状态信息可以直接写入目的服务器的微处理器和微码中；

由于 BAR 寄存器中包括 BAR 只读寄存器和 BAR 可写寄存器，从源端 SR-IOV 网卡 BAR 寄存器获取的该状态信息，恢复时需分两种情况进行处理，下面分别进行介绍：

1、从 BAR 可写寄存器中拷贝的该状态信息可以直接写入目的服务器的 BAR 可写寄存器；

2、从 BAR 只读寄存器中拷贝的该状态信息需要通过微处理器进行恢复，写入目的服务器 SR-IOV 网卡 BAR 只读寄存器中。

目的服务器的虚拟机管理器对该状态信息进行恢复后，可以使能网卡，进而启用第二虚拟机，完成该第一虚拟机的热迁移。

可选的，虚拟机管理器通过 VFIO 向热迁移模块发送控制命令。

本申请实施例中目的服务器接收源服务器发送的第一虚拟网卡的状态信息，可以通过新增的热迁移模块，实现该状态信息在 SR-IOV 网卡中的恢复，使得第一虚拟机的成功迁移。

基于图 2 提供的虚拟机热迁移的系统架构，请参阅图 5，为本申请实施例中虚拟机热迁移方法的一个实施例交互示意图。

5 虚拟机热迁移技术常用于保持服务器的负载均衡或虚拟机的容灾、备份等场景中。当源服务器中的第一虚拟机从源服务器迁移至目的服务器时，可以根据本申请实施例提供的虚拟机热迁移方案进行迁移。

501、源服务器与目的服务器建立连接；

10 当源服务器中的第一虚拟机从源服务器迁移至目的服务器时，源服务器获取热迁移请求，源服务器根据该热迁移请求与目的主机建立连接，并求向该目的服务器发送热迁移请求，指示源服务器中第一虚拟机迁移至该目的服务器，目的服务器可以进行热迁移前的准备工作，例如，目的服务器为该第一虚拟机预保留虚拟机资源。

可选的，源服务器接收用户的迁移请求，源服务器根据目的服务器的主机 IP 地址与目的主机建立连接。

15 502、目的服务器执行全局资源分配；

目的服务器执行全局资源分配，即初始化 VF 资源，保证 VF 资源的可用性。目的服务器中，虚拟机管理器通过热迁移模块在操作系统中查询并申请可用的全局资源，包括第二虚拟机对应的第二虚拟网卡的标识信息。

20 全局资源是虚拟机对应的虚拟网卡在操作系统中保存的虚拟网卡标识信息。例如索引号 (func_id)、RSS 模板表标识或 q_cfg 等。

示例性的，源服务器中，第一虚拟机对应的第一虚拟网卡的索引号是 1，迁移到目的服务器的第二虚拟机，目的服务器中，该索引号 1 可能被其他虚拟机占用，目的服务器为第二虚拟机对应的第二虚拟网卡分配的索引号可能是 2。

25 示例性的，VF 资源并不是独立分配的，例如 RSS 模板表，每个 VF 使用的 RSS 是微处理器全局分配的，在源端的 rss_id 是 1，但 1 在目的端可能已经分配给其他的 VF 了。

可选的，目的端虚拟机管理器通过 pre_active 接口进行全局资源的分配。

可选的，目的服务器中，虚拟机管理器分配虚拟网卡标识信息的步骤被称为 post_active 流程。

30 可选的，热迁移模块通过微处理器重新分配 VF 对应的全局资源，然后执行状态恢复流程。热迁移模块通过微处理器和微码设置芯片状态信息和 VM 对应资源的配置，流程返回后执行虚拟机 active 操作。本申请实施例提供的虚拟机热迁移方法，实现了迁移源端 VF 和迁移目的端目标 VF 的配置一致性，保证了虚拟机在迁移后的正常工作。

需要说明的是，步骤 502 在步骤 501 之后，步骤 508 之前执行，与步骤 503 至 507 的执行顺序不做限定。

35 503、源服务器迭代拷贝内存数据；

虚拟机管理器可以根据该热迁移请求，通过该热迁移模块执行内存数据迭代拷贝。第一虚拟机迁移内存数据的方式有多种，此处不做限定。通常，内存数据的拷贝需要执行多次，迭代拷贝至需要复制的内存小于预设门限后，将进入停机拷贝阶段，通常剩余的内存可一次

拷贝完成。

504、源服务器锁定并排空第一任务队列；

第一任务队列中用于放置为第一虚拟机下发的任务，该任务由用户通过应用层下发，该第一任务队列在实际应用过程中被称为执行用户命令的通道（command queue, cmdq）。

5 可选的，其中包括与第一虚拟机对应的虚拟网卡的状态信息相关的任务；

10 可选的，用户下发的任务为硬件操作，一次硬件操作可能是非原子性的，即涉及多次操作的组合。示例性的，VM 在热迁移时虚拟功能驱动可能正在操作硬件，然而完成一次硬件操作是多次寄存器或硬件交互的组合，该硬件操作为非原子性的，例如：VF 驱动向 PF 驱动发送 Mailbox 消息时，先写 Header 寄存器，再写数据寄存器，最后写 request 寄存器触发 VF 硬件开始读取头和数据发到目的 PF，PF 把处理结果返回给 VF 的异步事件队列（asynchronous event queue, AEQ），VFD 通过中断开始处理到 AEQ 中的结果。VM 迁移时被暂停在上述任何一个过程中，都会导致迁移到目的服务器时恢复出不同的状态，可能出现 Mailbox 消息丢失，或者发送了一个非法的 Mailbox 给目的端的 PF 等异常情况。

可选的，源服务器锁定并排空第一任务队列可称为 pre_deactivate 流程。

15 在第一虚拟机的迁移过程中，若用户通过该第一虚拟机下发用户命令改变对应的虚拟网卡的状态信息，将可能发生在命令执行完成之前，第一虚拟机对应的虚拟网卡状态信息已经迁移到目标虚拟机的情况，有可能产生命令超时等异常，导致业务中断。

20 为解决这一问题，本申请实施例中，热迁移模块将执行用户命令的通道锁定，以使得该通道无法接收用户下发的新任务。执行用户命令的通道被锁定后，用户将无法修改第一虚拟机对应的虚拟网卡的状态信息，但用户业务仍可持续进行。

该第一虚拟机运行的第一虚拟功能驱动排空执行用户命令的通道，是指将锁定队列时已经存在的用户命令执行完成，将队列排空，排空完成代表已下发的用户命令执行完成。热迁移模块可以获取该执行用户命令的通道排空的信息。

25 需要说明的是，执行用户命令的通道的状态属于待迁移的第一虚拟网卡的状态信息，源服务器将执行用户命令的通道锁定后，该锁定状态将迁移至目的服务器，目的服务器恢复该状态信息后，可以解锁该执行用户命令的通道。

为了进一步介绍源服务器锁定并排空第一任务队列的实现过程，请参阅图 6，本申请实施例中锁定并排空任务队列的一个实施例示意图。

一、热迁移模块侧：

30 611、热迁移模块设置任务队列为锁定状态。

热迁移模块可以在获取第一虚拟网卡的状态信息之前，设置任务队列为锁定状态，可选的，在热迁移模块执行最后一次内存拷贝前，设置任务队列为锁定状态。

需要说明的是，该任务队列的锁定状态可以迁移至目的服务器侧，并在目的服务器侧解除锁定状态，用于继续执行用户命令。

35 612、热迁移模块判断任务队列是否清空，若是，则执行步骤 613，若否，则重复执行步骤 612。

613、若热迁移模块判断任务队列清空，则执行其他操作。

若任务队列清空，则热迁移模块可执行其他操作，例如，上报虚拟机管理器，具体此处

不做限定。

二、虚拟功能驱动侧：

621、虚拟功能驱动获取用户下发的任务。

622、虚拟功能驱动判断任务队列是否处于锁定状态。

5 623、若虚拟功能驱动任务队列不处于锁定状态，则虚拟功能驱动将该任务加入任务队列。

624、若虚拟功能驱动任务队列处于锁定状态，则将该任务加入任务队列。

若热迁移模块设置任务队列为锁定状态，此后，该任务队列不能再接收新的任务，若用户下发了新的任务，则虚拟功能驱动将该任务置于等待通道中。当虚拟机迁移完成后，在目的服务器侧，任务队列的锁定状态解除后，等待通道中的任务可以被加入任务队列继续执行。

10 625、虚拟功能驱动执行任务队列中的任务。

626、虚拟功能驱动将任务队列清空。

在具体的实现过程中，可以通过设置 `busy_state` 参数、`own` 参数等实现，其中，热迁移模块锁定任务队列时，设置 `busy_state` 为 1，任务队列中不再加新任务，解除锁定时 `busy_state` 为 0，可以继续接收任务。虚拟功能驱动处理任务时 `own` 为 1，任务清空时 `own` 为 0，`func_ref` 为待执行命令数。

15 源端 VM pause 前，通过 `pre_deactivate` 流程确保 VFD 消息通道 `suspend`，`pre_deactivate` 返回则表示所有下发的消息均已处理完成，此时 VM 可 `pause`。目的端 VM `resume` 前，通过 PFD `clear busy_state`。然后当目的 VM `resume` 后，`suspend` 的通道在检测到通道不处于 `busy` 状态后恢复，继续下发源端 VFD 积压的消息。迁移源端流程中：
20 `Pre_deactivate` 流程中设置 `busy_state` 为 1，然后检查 `Own`，直到 `Own` 为 0 退出；在 VFD `Uninterrupt flow` 流程中，`busy_state` 状态为 1，则执行消息处理，直到 `func_ref` 为 0 后 `clear own bit`，然后一直等待 `busy_state` 状态变为 0。

虚拟功能驱动执行任务队列中的任务，直至任务队列中的任务被清空。热迁移模块可以获取该任务队列的任务是否清空。

25 需要说明的是，步骤 504 可以在步骤 501 至步骤 505 之间任一时刻执行，由于锁定执行用户命令的通道后，将不再响应新的用户命令。可选的，在步骤 503 之后，当迭代拷贝内存结束，剩余内存可一次拷贝完成时（`log_sync` 时刻）执行步骤 504，可以减少不响应用户命令阶段的时长。

505、源服务器运行的第一虚拟机停机；

30 源服务器确认执行用户命令的通道排空之后，虚拟机管理器调用虚拟机停机流程，挂起第一虚拟机。第一虚拟网卡停用，业务中断，第一虚拟机停机。

506、源服务器向目的服务器执行最后一次内存拷贝；

第一虚拟机停机之后，虚拟机管理器将执行最后一次内存拷贝，将步骤 503 中剩余的未拷贝内存迁移至目的服务器。

35 507、源服务器获取第一虚拟网卡的状态信息；

虚拟机管理器获取第一虚拟网卡的状态信息，并将该状态信息拷贝到目的服务器的虚拟机管理器中。

虚拟机管理器通过热迁移模块获取该第一虚拟机对应的第一虚拟网卡的状态信息。具体

地，热迁移模块与管理单元、数据处理单元协同工作，由虚拟机管理器分配的第一内存空间，热迁移模块根据状态信息存储位置分别从网卡的 BAR 寄存器、微处理器和微码中获取第一虚拟网卡的状态信息，并将该状态信息保存到该第一内存空间。需要说明的是，该第一内存空间不属于第一虚拟机对应的内存空间，由虚拟机管理器直接分配控制。

5 508、源服务器将虚拟网卡的状态信息发送至目的服务器进行恢复；

源服务器的虚拟机管理器将存储在该第一内存空间的该状态信息拷贝到目的服务器的虚拟机管理器。

目的服务器的虚拟机管理器接收到该状态信息后，需要将该状态信息分别恢复至 SR-IOV 网卡的 BAR 寄存器、微码和微处理器中。

10 从源服务器的微处理器和微码获取的该状态信息可以直接写入目的服务器的微处理器和微码中；

由于 BAR 寄存器中包括 BAR 只读寄存器和 BAR 可写寄存器，从源端 SR-IOV 网卡 BAR 寄存器获取的该状态信息，恢复时需分两种情况进行处理，下面分别进行介绍：

1、从 BAR 可写寄存器中拷贝的该状态信息可以直接写入目的服务器的 BAR 可写寄存器；

15 2、从 BAR 只读寄存器中拷贝的该状态信息需要通过微处理器进行恢复，写入目的服务器 SR-IOV 网卡 BAR 只读寄存器中。示例性的，例如报文发送通知 (Doorbell) 和配置控制通知 (Mailbox Req) 寄存器不能通过源端读取目的端写入的方式来恢复值，因此驻留在 BAR 只读寄存器的芯片状态不能直接借助 QEMU 拷贝 MMIO 空间的方式来迁移。

20 可选的，该虚拟机管理器接收所述源服务器发送的第二信息，该第二信息为第二虚拟机对应的第二虚拟网卡的队列的锁定状态信息，该第二虚拟网卡也是基于 SR-IOV 网卡虚拟出的，该第二任务队列用于放置为所述第二虚拟机下发的任务；所述热迁移模块解锁所述第二任务队列。

具体地，与步骤 504 对应地，目的服务器在启用第二虚拟机之前可将任务队列解锁，例如，通过设置 busy_state 为 0，将任务队列的锁定状态解锁。

25 509、目的服务器启用第二虚拟机；

第一虚拟网卡的状态信息迁移到目的服务器后，目的端虚拟机管理器将锁定的执行用户命令的通道打开，由此可以正常执行用户命令。

目的服务器的虚拟机管理器可以使能网卡，启用第二虚拟机。可选的，虚拟机管理器通过 activate 接口使能网卡。

30 510、目的服务器发送广播报文，恢复通信；

为缩短中断时延，目的服务器的虚拟机管理器发送广播报文，使得通信快速恢复。

可选的，虚拟机管理器通过 post_resume 接口发送广播报文。

需要说明的是，步骤 510 为可选步骤，可以执行，也可以不执行，此处不做限定。

35 本申请实施例中通过源服务器通过新增的热迁移模块，可以获取网卡中的第一虚拟网卡的状态信息，并将该状态信息拷贝至虚拟机管理器分配的内存中，进而发送至第二虚拟机。目的服务器接收源服务器发送的第一虚拟网卡的状态信息，可以通过新增的热迁移模块，实现该状态信息在 SR-IOV 网卡中的恢复，该虚拟机迁移系统可以实现第一虚拟网卡的状态信息迁移，使得第一虚拟机的成功迁移。

上面介绍了本申请实施例提供的虚拟机热迁移方法，下面对实现该方法的通信设备进行介绍，请参阅图 7，为本申请实施例中源端通信设备的一个实施例示意图。

本申请实施例提供了一种通信设备，可选的，该通信设备为虚拟机热迁移的源服务器，该源服务器中运行有第一虚拟机、该第一虚拟机使用的第一虚拟网卡、虚拟机管理器 701 和热迁移模块 702，该第一虚拟网卡是基于 SR-IOV 网卡虚拟出的，包括：

该虚拟机管理器 701，用于向该热迁移模块 702 下发热迁移命令；

该热迁移模块 702，用于根据该热迁移命令获取该第一虚拟机需迁移的信息，该第一虚拟机需迁移的信息包括该第一虚拟网卡的状态信息以及该第一虚拟机的内存数据，并上报该虚拟机管理器 701；

该虚拟机管理器 701，还用于向该目的服务器发送该状态信息以及该第一虚拟机的内存数据。

该热迁移模块 702 具体用于：根据该热迁移命令，将该状态信息拷贝至该虚拟机管理器 701 分配的该第一内存空间，该第一内存空间不包括在给该第一虚拟机分配的虚拟内存内。

该热迁移模块 702 具体用于：从该 SR-IOV 网卡的基地址寄存器、微处理器和微码中拷贝该状态信息至该第一内存空间。

该热迁移模块 702，还用于锁定该第一虚拟网卡的第一任务队列，以使得该第一任务队列无法接收新的任务，该第一任务队列用于放置为该第一虚拟机下发的任务；该服务器还包括：该第一虚拟机的第一虚拟功能驱动，用于清空该第一任务队列中所有已有的任务。

该热迁移模块 702 还用于：对该内存数据进行迭代拷贝；在剩余的该内存数据可一次拷贝完成的情况下，该第一虚拟功能驱动清空该第一任务队列中所有已有的任务之后，该虚拟机管理器 701 还用于，挂起该第一虚拟机；该热迁移模块 702，还用于执行最后一次内存数据拷贝。

请参阅图 8，为本申请实施例中目的端通信设备的一个实施例示意图。

本申请实施例提供一种通信设备，可选的，该通信设备为虚拟机热迁移的目的服务器，该目的服务器包括 SR-IOV 网卡，该目的服务器中运行有虚拟机管理器 801 和热迁移模块 802，该服务器包括：该虚拟机管理器 801，用于获取热迁移命令，该热迁移命令用于指示源服务器的第一虚拟机将迁移至该目的服务器的第二虚拟机；该虚拟机管理器 801，还用于接收该源服务器发送的该第一虚拟机对应的第一虚拟网卡的状态信息；该热迁移模块 802，还用于根据该虚拟机管理器 801 的指示，将该状态信息恢复至该 SR-IOV 网卡。

该热迁移模块 802 具体用于：将该状态信息分别恢复至该 SR-IOV 网卡的基地址寄存器、微码和微处理器中。

该热迁移模块 802 具体用于：将该状态信息中的第一信息拷贝至该微处理器，该第一信息为从该源服务器的只读基地址寄存器中获取的信息，该微处理器用于将该第一信息恢复至该 SR-IOV 网卡的只读基地址寄存器。

该虚拟机管理器 801 还用于：通过该热迁移模块 802 分配该第二虚拟机对应的第二虚拟网卡的标识信息。

该虚拟机管理器 801 具体用于：接收该源服务器发送的第二信息，该第二信息为该第二虚拟机对应的第二虚拟网卡的第二任务队列的锁定状态信息，该第二虚拟网卡是基于该 SR-

IOV 网卡虚拟出的，该第二任务队列用于放置为该第二虚拟机下发的任务；该热迁移模块 802，还用于解锁该第二任务队列。

请参阅图 9，为本申请实施例中一种通信设备的另一个实施例示意图。本实施例提供的通信设备为虚拟机热迁移的源服务器，本申请实施例中对其具体设备形态不做限定。

5 该源端通信设备 900 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上处理器 901 和存储器 902，该存储器 902 中存储有程序或数据。

其中，存储器 902 可以是易失性存储或非易失性存储。可选地，处理器 901 是一个或多个中央处理器（CPU，Central Processing Unit，该 CPU 可以是单核 CPU，也可以是多核 CPU。处理器 901 可以与存储器 902 通信，在源端通信设备 900 上执行存储器 902 中的一系列指令。

10 该源端通信设备 900 还包括物理网卡 903，该物理网卡 903 为 PCIE 设备，即 SR-IOV 网卡。

可选地，尽管图 9 中未示出，源端通信设备 900 还可以包括一个或一个以上电源；一个或一个以上有线或无线网络接口，例如以太网接口；一个或一个以上输入输出接口，输入输出接口可以用于连接显示器、鼠标、键盘、触摸屏设备或传感设备等，输入输出接口为可选
15 部件，可以存在也可以不存在，此处不做限定。

本实施例中源端通信设备 900 中的处理器 901 所执行的流程可以参考前述方法实施例中描述的方法流程，此处不加赘述。

请参阅图 10，为本申请实施例中一种通信设备的另一个实施例示意图。本实施例提供的通信设备为虚拟机热迁移的目的服务器，本申请实施例中对其具体设备形态不做限定。

20 该目的端通信设备 1000 可因配置或性能不同而产生比较大的差异，可以包括一个或一个以上处理器 1001 和存储器 1002，该存储器 1002 中存储有程序或数据。

其中，存储器 1002 可以是易失性存储或非易失性存储。可选地，处理器 1001 是一个或多个中央处理器（CPU，Central Processing Unit，该 CPU 可以是单核 CPU，也可以是多核 CPU。处理器 1001 可以与存储器 1002 通信，在目的端通信设备 1000 上执行存储器 1002 中的一系列指令。
25

该目的端通信设备 1000 还包括物理网卡 1003，该物理网卡 1003 为 PCIE 设备，即 SR-IOV 网卡。

可选地，尽管图 10 中未示出，目的端通信设备 1000 还可以包括一个或一个以上电源；一个或一个以上有线或无线网络接口，例如以太网接口；一个或一个以上输入输出接口，输入输出接口可以用于连接显示器、鼠标、键盘、触摸屏设备或传感设备等，输入输出接口为
30 可选部件，可以存在也可以不存在，此处不做限定。

本实施例中目的端通信设备 1000 中的处理器 1001 所执行的流程可以参考前述方法实施例中描述的方法流程，此处不加赘述。

所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统，装置
35 和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统，装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结

合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本申请各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用软件功能单元的形式实现。

所述集成的单元如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中。基于这样的理解，本申请的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的全部或部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备）执行本申请各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U 盘、移动硬盘、只读存储器（read-only memory，ROM）、随机存取存储器（random access memory，RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

以上所述，以上实施例仅用以说明本申请的技术方案，而非对其限制；尽管参照前述实施例对本申请进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本申请各实施例技术方案的精神和范围。

权利要求

1、一种虚拟机热迁移方法，其特征在于，应用于运行有第一虚拟机的源服务器，所述方法用于向目的服务器迁移所述第一虚拟机，所述源服务器包括单根输入/输出虚拟化 SR-IOV 网卡，所述源服务器中运行有第一虚拟机、所述第一虚拟机使用的第一虚拟网卡、虚拟机管理器 5 和热迁移模块，所述第一虚拟网卡是基于所述 SR-IOV 网卡虚拟出的，所述方法包括：

所述虚拟机管理器向所述热迁移模块下发热迁移命令；

所述热迁移模块根据所述热迁移命令获取所述第一虚拟机需迁移的信息，所述第一虚拟机需迁移的信息包括所述第一虚拟网卡的状态信息以及所述第一虚拟机的内存数据，并上报所述虚拟机管理器；

10 所述虚拟机管理器向所述目的服务器发送所述状态信息以及所述第一虚拟机的内存数据。

2、根据权利要求 1 所述的方法，其特征在于，所述热迁移模块根据所述热迁移命令获取所述第一虚拟机需迁移的信息包括：

所述热迁移模块根据所述热迁移命令，将所述状态信息拷贝至所述虚拟机管理器分配的所述第一内存空间，所述第一内存空间不包括在给所述第一虚拟机分配的虚拟内存内。

15 3、根据权利要求 2 所述的方法，其特征在于，所述热迁移模块根据所述热迁移命令，将所述状态信息拷贝至所述虚拟机管理器分配的所述第一内存空间包括：

所述热迁移模块从所述 SR-IOV 网卡的基地址寄存器、微处理器和微码中拷贝所述状态信息至所述第一内存空间。

4、根据权利要求 1 至 3 中任一项所述的方法，其特征在于，所述热迁移模块根据所述热迁移命令获取所述第一虚拟机需迁移的信息之前，所述方法还包括：

所述热迁移模块锁定所述第一虚拟网卡的第一任务队列，以使得所述第一任务队列无法接收新的任务，所述第一任务队列用于放置为所述第一虚拟机下发的任务；

所述第一虚拟机的第一虚拟功能驱动清空所述第一任务队列中所有已有的任务。

25 5、根据权利要求 4 所述的方法，其特征在于，所述虚拟机管理器向所述热迁移模块下发热迁移命令之后，所述方法还包括：

所述热迁移模块对所述内存数据进行迭代拷贝；

在剩余的所述内存数据可一次拷贝完成的情况下，所述第一虚拟功能驱动清空所述第一任务队列中所有已有的任务之后，所述虚拟机管理器挂起所述第一虚拟机；

所述热迁移模块执行最后一次内存数据拷贝。

30 6、一种虚拟机热迁移方法，其特征在于，应用于虚拟机迁移的目的服务器，所述目的服务器包括 SR-IOV 网卡，所述目的服务器中运行有虚拟机管理器和热迁移模块，所述方法包括：

所述虚拟机管理器获取热迁移命令，所述热迁移命令用于指示源服务器的第一虚拟机将迁移至所述目的服务器的第二虚拟机；

35 所述虚拟机管理器接收所述源服务器发送的所述第一虚拟机对应的第一虚拟网卡的状态信息；

所述热迁移模块根据所述虚拟机管理器的指示，将所述状态信息恢复至所述 SR-IOV 网卡。

7、根据权利要求 6 所述的方法，其特征在于，所述热迁移模块根据所述虚拟机管理器的指示，将所述状态信息恢复至所述 SR-IOV 网卡包括：

所述热迁移模块将所述状态信息分别恢复至所述 SR-IOV 网卡的基地址寄存器、微码和微处理器中。

8、根据权利要求 7 所述的方法，其特征在于，所述热迁移模块将所述状态信息分别恢复至所述 SR-IOV 网卡的基地址寄存器、微码和微处理器中包括：

5 所述热迁移模块将所述状态信息中的第一信息拷贝至所述微处理器，所述第一信息为从所述源服务器的只读基地址寄存器中获取的信息，所述微处理器用于将所述第一信息恢复至所述 SR-IOV 网卡的只读基地址寄存器。

9、根据权利要求 6 至 8 中任一项所述的方法，其特征在于，所述虚拟机管理器接收所述源服务器发送的所述第一虚拟机对应的第一虚拟网卡的状态信息之前，所述方法还包括：

10 所述虚拟机管理器通过所述热迁移模块分配所述第二虚拟机对应的第二虚拟网卡的标识信息。

10、根据权利要求 6 至 9 中任一项所述的方法，其特征在于，所述虚拟机管理器接收所述源服务器发送的所述第一虚拟机对应的第一虚拟网卡的状态信息包括：

15 所述虚拟机管理器接收所述源服务器发送的第二信息，所述第二信息为所述第二虚拟机对应的第二虚拟网卡的第二任务队列的锁定状态信息，所述第二虚拟网卡是基于所述 SR-IOV 网卡虚拟出的，所述第二任务队列用于放置为所述第二虚拟机下发的任务；

所述方法还包括：所述热迁移模块解锁所述第二任务队列。

11、一种通信设备，其特征在于，该通信设备为虚拟机热迁移的源服务器，所述源服务器中运行有第一虚拟机、所述第一虚拟机使用的第一虚拟网卡、虚拟机管理器和热迁移模块，
20 所述第一虚拟网卡是基于 SR-IOV 网卡虚拟出的，所述通信设备包括：

所述虚拟机管理器，用于向所述热迁移模块下发热迁移命令；

所述热迁移模块，用于根据所述热迁移命令获取所述第一虚拟机需迁移的信息，所述第一虚拟机需迁移的信息包括所述第一虚拟网卡的状态信息以及所述第一虚拟机的内存数据，
并上报所述虚拟机管理器；

25 所述虚拟机管理器，还用于向所述目的服务器发送所述状态信息以及所述第一虚拟机的内存数据。

12、根据权利要求 11 所述的通信设备，其特征在于，所述热迁移模块具体用于：

根据所述热迁移命令，将所述状态信息拷贝至所述虚拟机管理器分配的所述第一内存空间，所述第一内存空间不包括在给所述第一虚拟机分配的虚拟内存内。

30 13、根据权利要求 12 所述的通信设备，其特征在于，所述热迁移模块具体用于：

从所述 SR-IOV 网卡的基地址寄存器、微处理器和微码中拷贝所述状态信息至所述第一内存空间。

14、根据权利要求 11 至 13 中任一项所述的通信设备，其特征在于，

35 所述热迁移模块，还用于锁定所述第一虚拟网卡的第一任务队列，以使得所述第一任务队列无法接收新的任务，所述第一任务队列用于放置为所述第一虚拟机下发的任务；

所述通信设备还包括：

所述第一虚拟机的第一虚拟功能驱动，用于清空所述第一任务队列中所有已有的任务。

15、根据权利要求 14 所述的通信设备，其特征在于，所述热迁移模块还用于：

对所述内存数据进行迭代拷贝；

在剩余的所述内存数据可一次拷贝完成的情况下，所述第一虚拟功能驱动清空所述第一任务队列中所有已有的任务之后，所述虚拟机管理器还用于，挂起所述第一虚拟机；

所述热迁移模块，还用于执行最后一次内存数据拷贝。

5 16、一种通信设备，其特征在于，该通信设备为虚拟机热迁移的目的服务器，所述目的服务器包括 SR-IOV 网卡，所述目的服务器中运行有虚拟机管理器和热迁移模块，所述通信设备包括：

所述虚拟机管理器，用于获取热迁移命令，所述热迁移命令用于指示源服务器的第一虚拟机将迁移至所述目的服务器的第二虚拟机；

10 所述虚拟机管理器，还用于接收所述源服务器发送的所述第一虚拟机对应的第一虚拟网卡的状态信息；

所述热迁移模块，还用于根据所述虚拟机管理器的指示，将所述状态信息恢复至所述 SR-IOV 网卡。

15 17、根据权利要求 16 所述的通信设备，其特征在于，所述热迁移模块具体用于：
将所述状态信息分别恢复至所述 SR-IOV 网卡的基地址寄存器、微码和微处理器中。

18、根据权利要求 17 所述的通信设备，其特征在于，所述热迁移模块具体用于：

将所述状态信息中的第一信息拷贝至所述微处理器，所述第一信息为从所述源服务器的只读基地址寄存器中获取的信息，所述微处理器用于将所述第一信息恢复至所述 SR-IOV 网卡的只读基地址寄存器。

20 19、根据权利要求 16 至 18 中任一项所述的通信设备，其特征在于，所述虚拟机管理器还用于：

通过所述热迁移模块分配所述第二虚拟机对应的第二虚拟网卡的标识信息。

20、根据权利要求 16 至 19 中任一项所述的通信设备，其特征在于，所述虚拟机管理器具体用于：

25 接收所述源服务器发送的第二信息，所述第二信息为所述第二虚拟机对应的第二虚拟网卡的第二任务队列的锁定状态信息，所述第二虚拟网卡是基于所述 SR-IOV 网卡虚拟出的，所述第二任务队列用于放置为所述第二虚拟机下发的任务；

所述热迁移模块，还用于解锁所述第二任务队列。

21、一种通信设备，包括：处理器、存储器和 SR-IOV 网卡；

30 所述处理器用于基于所述存储器中的代码运行第一虚拟机、虚拟机管理器和热迁移模块，所述第一虚拟机包括第一虚拟功能驱动，所述处理器通过所述虚拟机管理器、所述热迁移模块和所述第一虚拟功能驱动执行如权利要求 1 至 5 中任一项所述的方法。

22、一种通信设备，包括：处理器、存储器和 SR-IOV 网卡；

35 所述处理器用于基于所述存储器中的代码运行虚拟机管理器和热迁移模块，并通过所述虚拟机管理器和热迁移模块执行如权利要求 6 至 10 中任一项所述的方法。

23、一种包含指令的计算机程序产品，其特征在于，当其在计算机上运行时，使得所述计算机执行如权利要求 1 至 10 任一项所述的方法。

24、一种计算机可读存储介质，包括指令，其特征在于，当所述指令在计算机上运行时，

使得计算机执行如权利要求 1 至 10 中任一项所述的方法。

25、一种虚拟机热迁移系统，包括如权利要求 11 至 15 中任一项所述源服务器，和如权利要求 16 至 20 中任一项所述目的服务器。

5

10

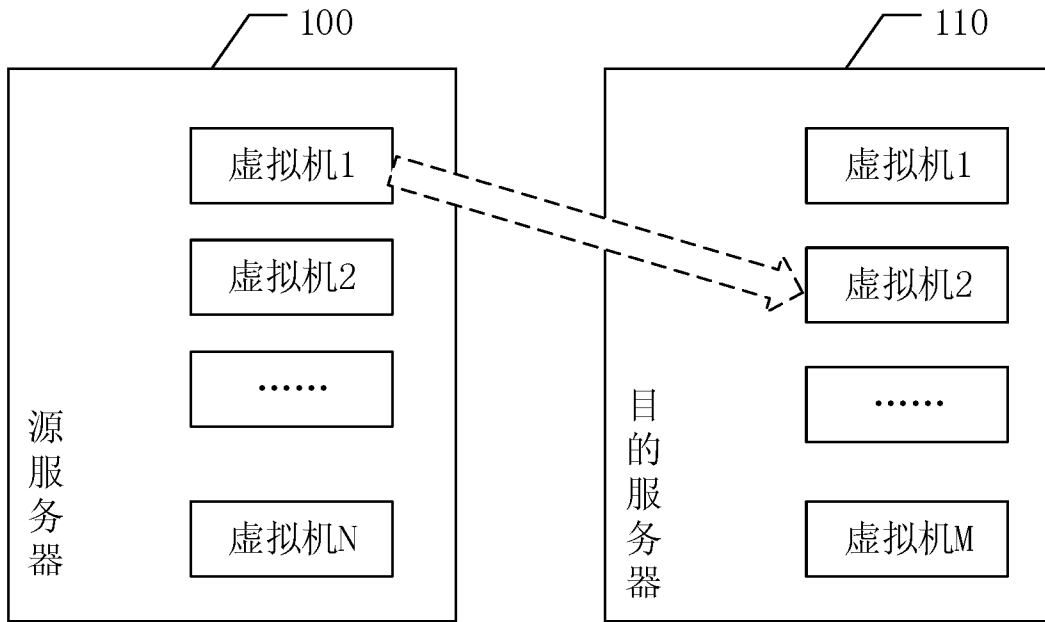


图 1

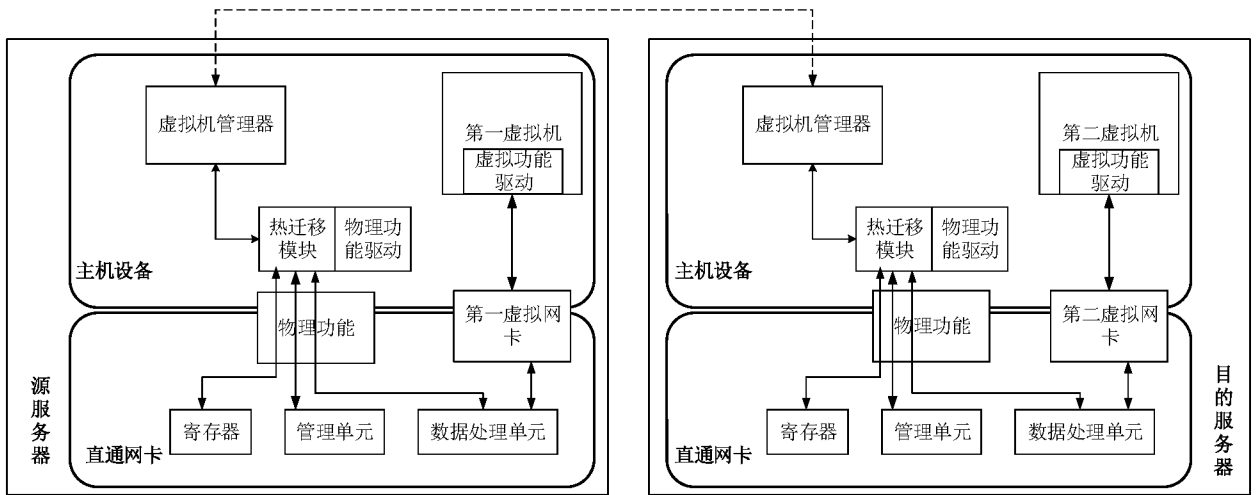


图 2

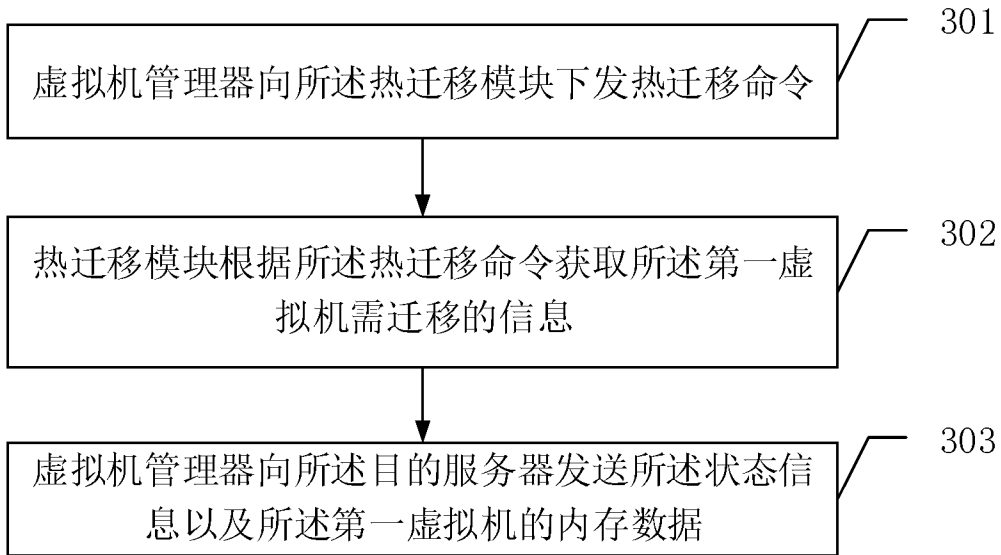


图 3

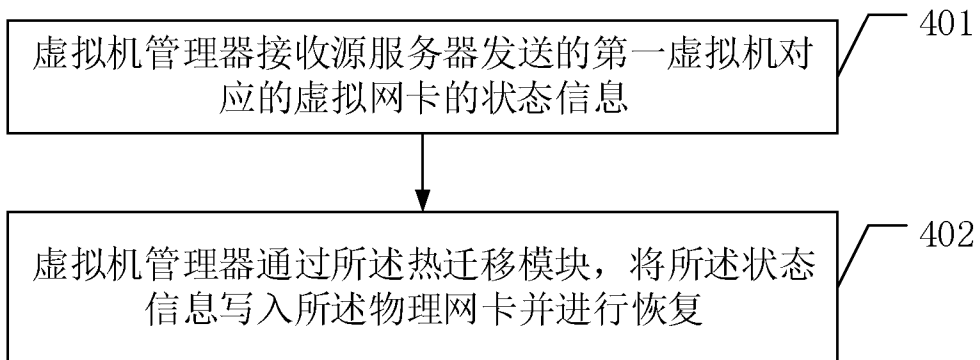


图 4

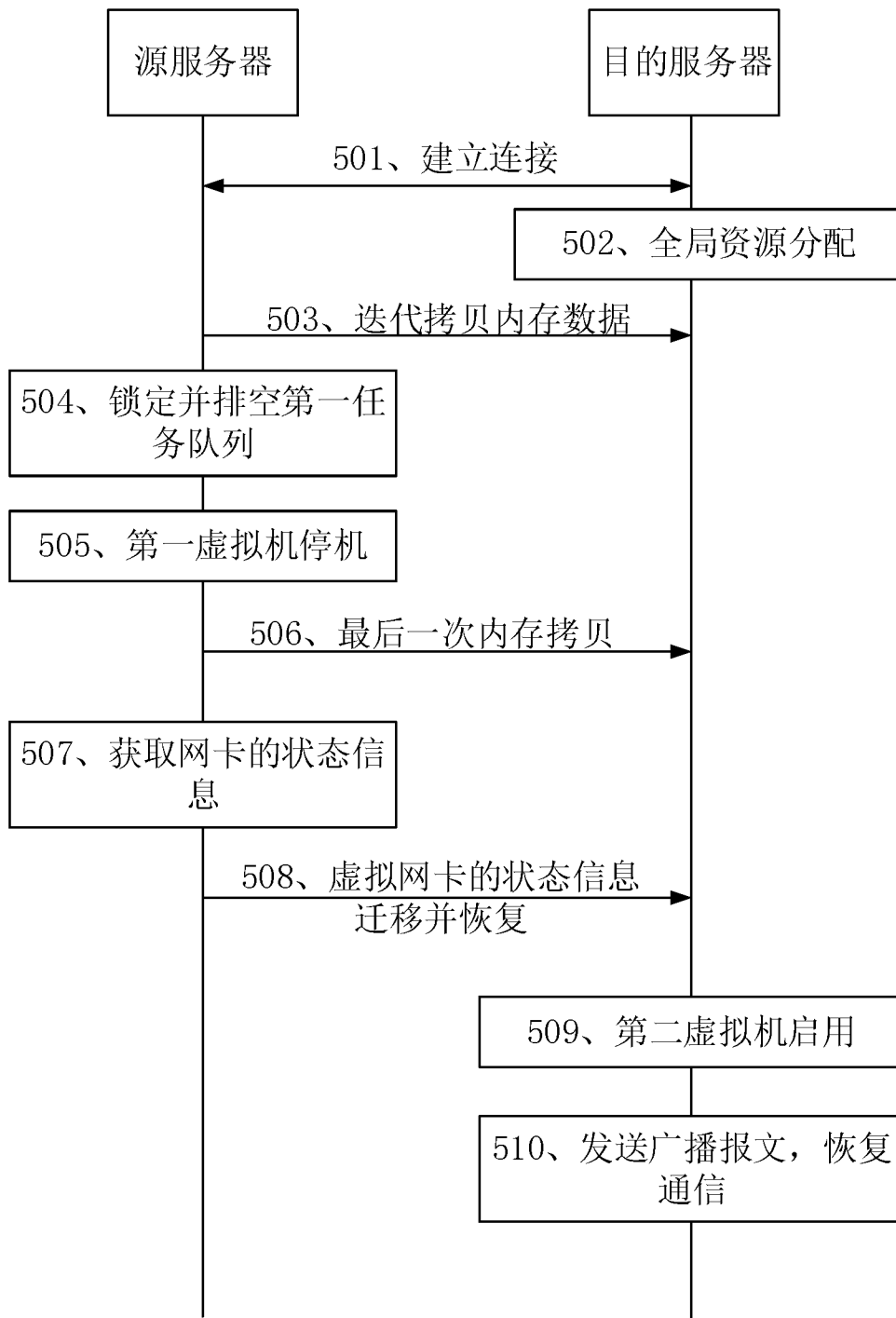


图 5

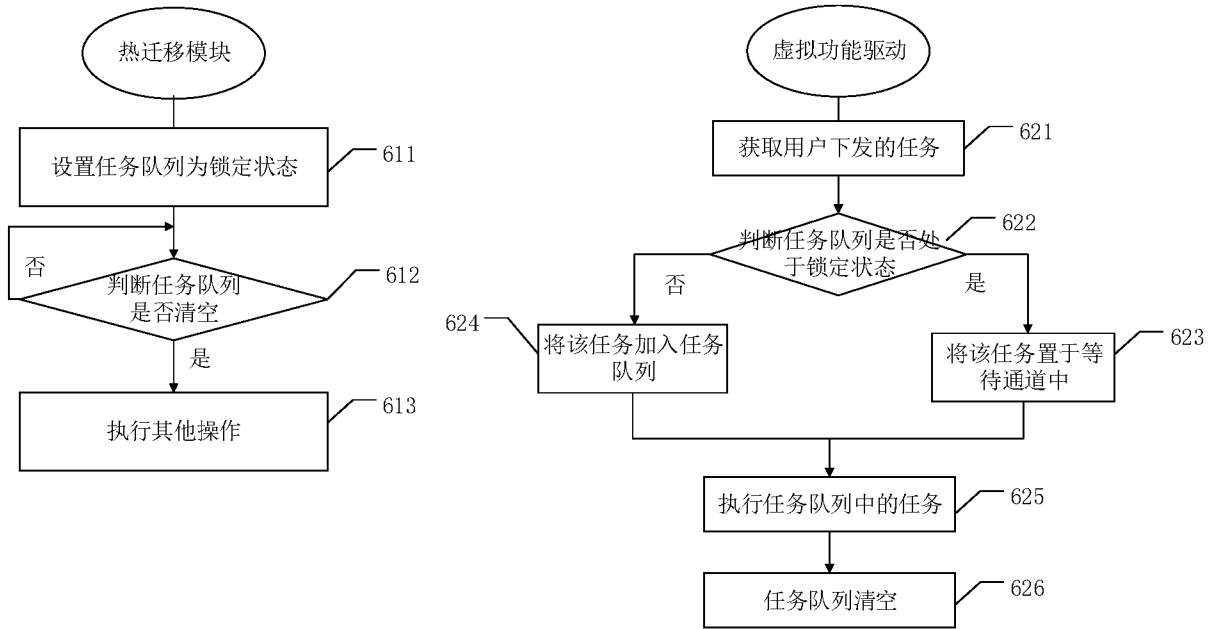


图 6

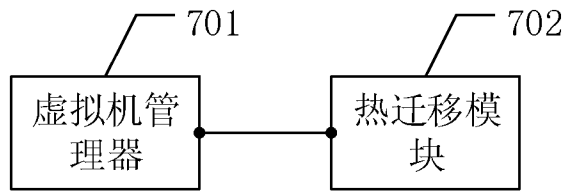


图 7

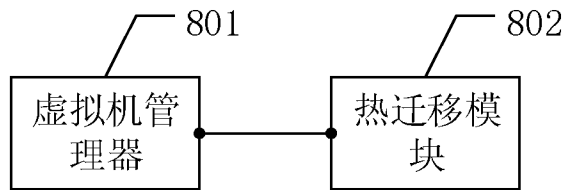


图 8

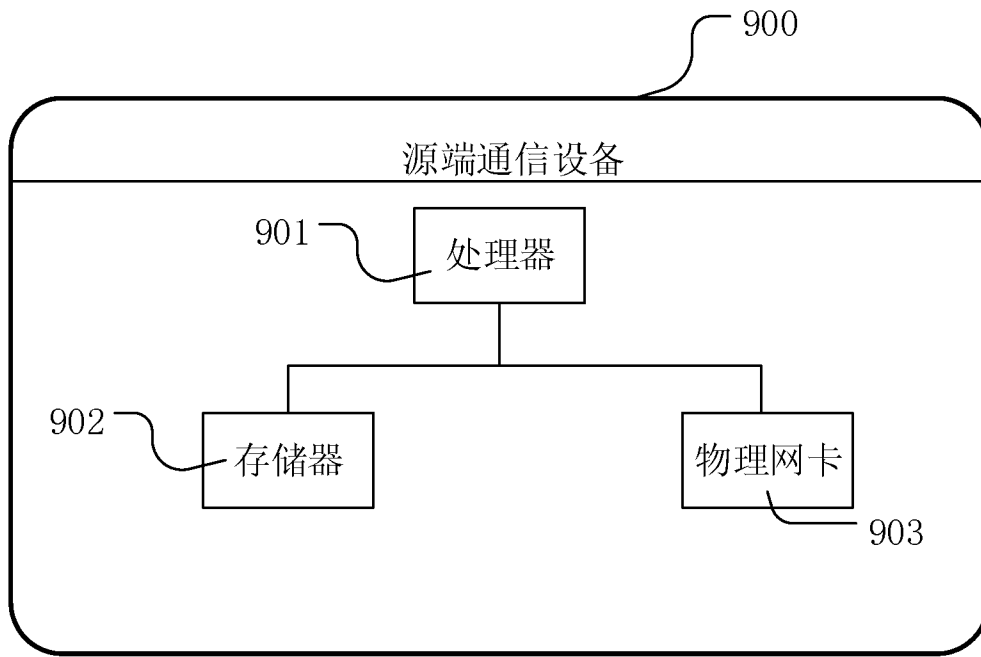


图 9

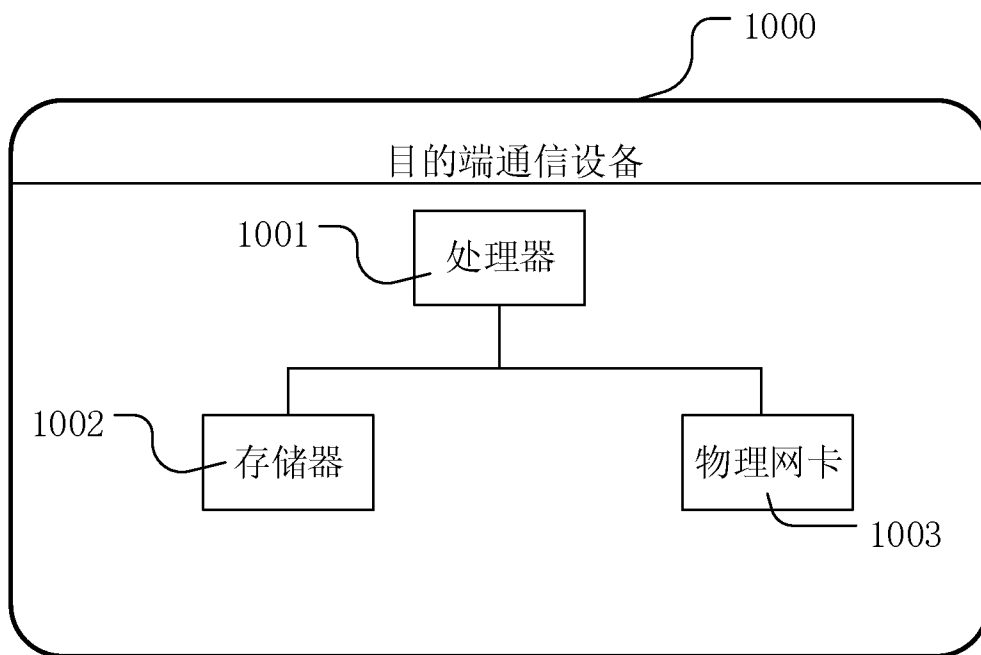


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/106044

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) 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) CNTXT; CNABS; CNKI; WOTXT; USTXT; VEN; EPTXT: 拷贝, 热, 迁移, 单根输入输出, 内存, 状态, 虚拟机, 管理, 锁, 新任务, 网卡, copy, heat, migrate, single root I/O, memory, condition, virtual machine, manage, lock, new task, network card		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 108205506 A (HUAWEI TECHNOLOGIES CO., LTD.) 26 June 2018 (2018-06-26) claims 1-14, description paragraphs 2-7, 194-358	1, 4-11, 14-25
Y	CN 103955399 A (HUAWEI TECHNOLOGIES CO., LTD.) 30 July 2014 (2014-07-30) description, paragraphs 48-75	1, 4-11, 14-25
Y	CN 106464579 A (GOOGLE INC.) 22 February 2017 (2017-02-22) description, paragraph 33	4-5, 10, 14-15, 20-25
A	US 2012011508 A1 (VMWARE INC.) 12 January 2012 (2012-01-12) entire document	1-25
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 22 October 2020		Date of mailing of the international search report 02 November 2020
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/CN) No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088 China Facsimile No. (86-10)62019451		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/106044

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	108205506	A	26 June 2018	None			
CN	103955399	A	30 July 2014	CN	103955399	B	20 June 2017
CN	106464579	A	22 February 2017	US	2018203721	A1	19 July 2018
				US	10013276	B2	03 July 2018
				JP	2017518568	A	06 July 2017
				AU	2018201833	A1	12 April 2018
				AU	2015277109	A1	24 November 2016
				KR	101862273	B1	29 May 2018
				EP	3135007	A1	01 March 2017
				WO	2015195913	A1	23 December 2015
				US	10592273	B2	17 March 2020
				AU	2018201833	B2	21 February 2019
				CN	106464579	B	03 December 2019
				KR	20160146984	A	21 December 2016
				EP	3135007	B1	06 May 2020
				JP	6534399	B2	26 June 2019
				AU	2015277109	B2	01 March 2018
				US	2015370596	A1	24 December 2015
				WO	2015195913	A8	01 December 2016
				IN	201647038360	A	09 December 2016
US	2012011508	A1	12 January 2012	US	8990531	B2	24 March 2015

国际检索报告

国际申请号

PCT/CN2020/106044

<p>A. 主题的分类</p> <p>G06F 13/28 (2006.01) i</p> <p>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</p>																	
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>G06F</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CNXTX;CNABS;CNKI;WOTXT;USTXT;VEN;EPTXT;拷贝, 热, 迁移, 单根输入输出, 内存, 状态, 虚拟机, 管理, 锁, 新任务, 网卡,copy, heat, migrate, single root I/O, memory, condition, virtual machine, manage, lock, new task, network card</p>																	
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>CN 108205506 A (华为技术有限公司) 2018年 6月 26日 (2018 - 06 - 26) 权利要求1-14, 说明书第2-7、194-358段</td> <td>1、4-11、14-25</td> </tr> <tr> <td>Y</td> <td>CN 103955399 A (华为技术有限公司) 2014年 7月 30日 (2014 - 07 - 30) 说明书第48-75段</td> <td>1、4-11、14-25</td> </tr> <tr> <td>Y</td> <td>CN 106464579 A (谷歌公司) 2017年 2月 22日 (2017 - 02 - 22) 说明书第33段</td> <td>4-5、10、 14-15、20-25</td> </tr> <tr> <td>A</td> <td>US 2012011508 A1 (VMWARE INC) 2012年 1月 12日 (2012 - 01 - 12) 全文</td> <td>1-25</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	CN 108205506 A (华为技术有限公司) 2018年 6月 26日 (2018 - 06 - 26) 权利要求1-14, 说明书第2-7、194-358段	1、4-11、14-25	Y	CN 103955399 A (华为技术有限公司) 2014年 7月 30日 (2014 - 07 - 30) 说明书第48-75段	1、4-11、14-25	Y	CN 106464579 A (谷歌公司) 2017年 2月 22日 (2017 - 02 - 22) 说明书第33段	4-5、10、 14-15、20-25	A	US 2012011508 A1 (VMWARE INC) 2012年 1月 12日 (2012 - 01 - 12) 全文	1-25
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求															
Y	CN 108205506 A (华为技术有限公司) 2018年 6月 26日 (2018 - 06 - 26) 权利要求1-14, 说明书第2-7、194-358段	1、4-11、14-25															
Y	CN 103955399 A (华为技术有限公司) 2014年 7月 30日 (2014 - 07 - 30) 说明书第48-75段	1、4-11、14-25															
Y	CN 106464579 A (谷歌公司) 2017年 2月 22日 (2017 - 02 - 22) 说明书第33段	4-5、10、 14-15、20-25															
A	US 2012011508 A1 (VMWARE INC) 2012年 1月 12日 (2012 - 01 - 12) 全文	1-25															
<p><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。</p>																	
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&” 同族专利的文件</p>																	
<p>国际检索实际完成的日期</p> <p>2020年 10月 22日</p>		<p>国际检索报告邮寄日期</p> <p>2020年 11月 2日</p>															
<p>ISA/CN的名称和邮寄地址</p> <p>中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088</p> <p>传真号 (86-10)62019451</p>		<p>授权官员</p> <p>刘细金</p> <p>电话号码 86-(20)-28950395</p>															

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/106044

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	108205506	A	2018年 6月 26日	无			
CN	103955399	A	2014年 7月 30日	CN	103955399	B	2017年 6月 20日
CN	106464579	A	2017年 2月 22日	US	2018203721	A1	2018年 7月 19日
				US	10013276	B2	2018年 7月 3日
				JP	2017518568	A	2017年 7月 6日
				AU	2018201833	A1	2018年 4月 12日
				AU	2015277109	A1	2016年 11月 24日
				KR	101862273	B1	2018年 5月 29日
				EP	3135007	A1	2017年 3月 1日
				WO	2015195913	A1	2015年 12月 23日
				US	10592273	B2	2020年 3月 17日
				AU	2018201833	B2	2019年 2月 21日
				CN	106464579	B	2019年 12月 3日
				KR	20160146984	A	2016年 12月 21日
				EP	3135007	B1	2020年 5月 6日
				JP	6534399	B2	2019年 6月 26日
				AU	2015277109	B2	2018年 3月 1日
				US	2015370596	A1	2015年 12月 24日
				WO	2015195913	A8	2016年 12月 1日
				IN	201647038360	A	2016年 12月 9日
US	2012011508	A1	2012年 1月 12日	US	8990531	B2	2015年 3月 24日