

(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
2016年11月24日 (24.11.2016)



(10) 国际公布号
WO 2016/184352 A1

- (51) 国际专利分类号:
G06F 9/445 (2006.01)
- (21) 国际申请号: PCT/CN2016/081953
- (22) 国际申请日: 2016年5月13日 (13.05.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510264571.0 2015年5月21日 (21.05.2015) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛英属开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 张献涛 (ZHANG, Xiantao); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。 傅俊康 (FU, Junkang); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

- (74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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)。

[见续页]

(54) Title: METHOD AND DEVICE FOR UPGRADING VHOST-NET MODULE

(54) 发明名称: 一种半虚拟化网卡内核加速模块升级方法和装置

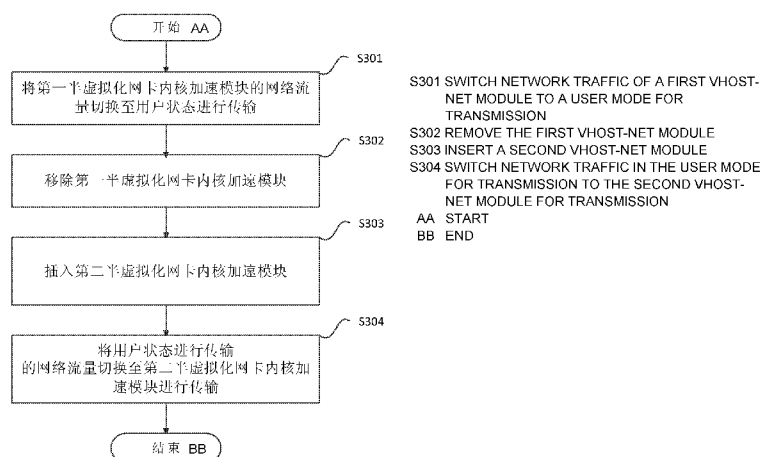


图 3

(57) Abstract: Provided are a method and device for upgrading a vhost-net module. The method comprises: switching network traffic of a first vhost-net module to a user mode and performing transmission; removing the first vhost-net module; inserting a second vhost-net module; and switching network traffic in the user mode for transmission to the second vhost-net module and performing transmission. With the method and device for upgrading the vhost-net module of the present invention, switching network traffic between a kernel mode and the user mode will not cause fluctuation in network traffic of a virtual machine (VM), and thus the upgrade of the vhost-net module is not noticeable to a user and will not affect user experience.

(57) 摘要: 本申请的目的是提供一种半虚拟化网卡内核加速模块升级方法和装置, 所述方法包括: 将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输; 移除第一半虚拟化网卡内核加速模块; 插入第二半虚拟化网卡内核加速模块; 将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输。本申请的半虚拟化网卡内核加速模块升级方法和装置, 网络流量在内核态和用户态之间的切换不会造成虚拟机网络流量的波动, 实现半虚拟化网卡内核加速模块的升级对于用户而言是无法感知的, 不会影响用户体验。



WO 2016/184352 A1

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

一种半虚拟化网卡内核加速模块升级方法和装置

本申请要求 2015 年 05 月 21 日递交的申请号为 201510264571.0、发明名称为“一种半虚拟化网卡内核加速模块升级方法和装置”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。

5

技术领域

本申请涉及计算机领域，尤其涉及一种半虚拟化网卡内核加速模块升级方法和装置。

背景技术

10 半虚拟化网卡内核加速模块（vhost-net）是基于内核的虚拟机（KVM，Kernel-based Virtual Machine）的重要组件，vhost-net 模块可以提升 KVM 虚拟机网络约 30% 的性能。

现有技术中，对于 vhost-net 模块升级时，需要将虚拟机关机，移除原 vhost-net 模块后，插入新的 vhost-net 模块，再重新启动虚拟机。这种 vhost-net 模块的升级方式必须暂时关闭虚拟机，关闭期间会造成服务中断，对于服务提供商和客户都会带来不良影响。

15

发明内容

本申请的一个目的是提供一种半虚拟化网卡内核加速模块升级方法和装置。

根据本申请的一个方面，提供了一种半虚拟化网卡内核加速模块升级方法，包括：

20 将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输；

移除第一半虚拟化网卡内核加速模块；

插入第二半虚拟化网卡内核加速模块；

将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输。

根据本申请的另一个方面，提供了一种半虚拟化网卡内核加速模块升级装置，包括：

25 第一模块，用于将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输；

第二模块，用于移除第一半虚拟化网卡内核加速模块；

第三模块，用于插入第二半虚拟化网卡内核加速模块；

30 第四模块，用于将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输。

与现有技术相比，本申请在对半虚拟化网卡内核加速模块进行升级时，首先将网络流量由内核态切换至用户态进行传输，然后将升级后的半虚拟化网卡内核加速模块插入系统，再将原先由用户态转发的网络流量切换回内核态由新的半虚拟化网卡内核加速模块进行转发，不需要中断服务，网络流量在内核态和用户态之间的切换不会造成虚拟机网络流量的波动，实现半虚拟化网卡内核加速模块的升级对于用户而言是无法感知的，不会影响用户体验。

附图说明

通过阅读参照以下附图所作的对非限制性实施例所作的详细描述，本申请的其它特征、目的和优点将会变得更明显：

图 1 示出根据本申请一个方面的半虚拟化网卡内核加速模块升级装置的结构示意图；

图 2 示出根据本申请又一个方面的半虚拟化网卡内核加速模块升级装置的结构示意图；

图 3 示出根据本申请一个方面的半虚拟化网卡内核加速模块升级方法的流程示意图；

图 4 示出根据本申请又一个方面的半虚拟化网卡内核加速模块升级方法的流程示意图。

附图中相同或相似的附图标记代表相同或相似的部件。

具体实施方式

下面结合附图对本申请作进一步详细描述。

图 1 示出根据本申请一个方面的半虚拟化网卡内核加速模块升级装置的结构示意图。如图 1 所示，该半虚拟化网卡内核加速模块升级装置 1 包括：第一模块 11、第二模块 12、第三模块 13 和第四模块 14。

其中，第一模块 11，用于将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输；第二模块 12，用于移除第一半虚拟化网卡内核加速模块；第三模块 13，用于插入第二半虚拟化网卡内核加速模块；第四模块 14，用于将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输。

在此，半虚拟化网卡内核加速模块可以采用 KVM 虚拟化的 vhost-net 模块，vhost-net

模块用于机器虚拟机（QEMU）半虚拟化网卡（virtio-net）加速。第一半虚拟化网卡内核加速模块是指升级前的 vhost-net 模块，第二半虚拟化网卡内核加速模块是指升级后新的 vhost-net 模块。第一模块 11 将旧的 vhost-net 模块的网络流量切换至用户态进行传输。在用户态下，虚拟机在发送或接收数据包时，首先陷入内核态，然后再返回到用户态的 QEMU（机器虚拟机）进程进行处理，最后通过虚拟以太网网络设备（TAP 设备）经由网桥、物理网卡进行网络数据包的发送或接收。在将旧的 vhost-net 模块的网络流量切换至用户态进行转发后，第二模块 12 可以移除旧的 vhost-net 模块，而网络服务不被中断。此后第三模块 13 插入新的 vhost-net 模块。第四模块 14 将用户态进行转发的网络流量切换至内核态由升级后新的 vhost-net 模块进行传输。内核态下，虚拟机陷入内核态后，不需要再返回到用户态，而由新的 vhost-net 模块的内核线程进行处理，通过虚拟以太网网络设备（TAP 设备）经由网桥、物理网卡进行数据包的发送或接收。通过用户态和内核态的互相切换，从而实现了在移除旧的 vhost-net 模块并插入新的升级后的 vhost-net 模块时，不会中断网络服务。

图 2 示出根据本申请又一个方面的半虚拟化网卡内核加速模块升级装置的结构示意图。如图 2 所示，该半虚拟化网卡内核加速模块升级装置 1 包括：第一模块 11、第二模块 12、第三模块 13 和第四模块 14。

其中，第一模块 11，用于将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输；第二模块 12，用于移除第一半虚拟化网卡内核加速模块；第三模块 13，用于插入第二半虚拟化网卡内核加速模块；第四模块 14，用于将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输。

在此，半虚拟化网卡内核加速模块可以采用 KVM 虚拟化的 vhost-net 模块，vhost-net 模块用于机器虚拟机（QEMU）半虚拟化网卡（virtio-net）加速。第一半虚拟化网卡内核加速模块是指升级前旧的 vhost-net 模块，第二半虚拟化网卡内核加速模块是指升级后新的 vhost-net 模块。第一模块 11 将旧的 vhost-net 模块的网络流量切换至用户态进行传输。在用户态下，虚拟机在发送或接收数据包时，首先陷入内核态，然后再返回到用户态的 QEMU 进程进行处理，最后通过虚拟以太网网络设备（TAP 设备）经由网桥、物理网卡进行数据包的发送或接收。在将旧的 vhost-net 模块的网络流量切换至用户态进行转发后，第二模块 12 可以移除旧的 vhost-net 模块，而网络服务不被中断。此后第三模块 13 插入新的 vhost-net 模块。第四模块 14 将用户态进行转发的网络流量切换至内核态由升级后新的 vhost-net 模块进行传输。内核态下，虚拟机陷入内核态后，不需要再返回到

用户态，而由新的 vhost-net 模块的内核线程进行处理，通过 TAP 设备（虚拟以太网网络设备）经由网桥、物理网卡进行数据包的发送或接收。通过用户态和内核态的互相切换，从而实现了在移除旧的 vhost-net 模块并插入新的升级后的 vhost-net 模块时，不会中断网络服务。

5 具体地，第一模块 11 包括：第一一子模块 111、第一二子模块 112 和第一三子模块 113。

其中，第一一子模块 111，用于机器虚拟机进程获取共享内存的使用状态。第一二子模块 112，用于所述机器虚拟机进程控制对共享内存中网络数据包的发送和/或接收。第一三子模块 113，用于停止所述第一半虚拟化网卡内核加速模块内核线程的轮询。

10 在此，第一一子模块 111 使 QEMU 进程获取共享内存（vring）的使用状态，第一二子模块 112 使 QEMU 进程控制对共享内存中网络数据包的发送和/或接收。第一三子模块 113 停止旧的 vhost-net 模块内核线程在 TAP 设备上的收/发网络数据包的轮询（poll），将旧的 vhost-net 模块的网络流量切换至用户态进行处理，为移除旧的 vhost-net 模块做好准备。其中，共享内存用于缓存收/发网络数据包，每个半虚拟化网卡对应一个 vhost-net
15 模块内核线程。

可选地，第二模块 12 还用于：清理第一半虚拟化网卡内核加速模块的资源。

在此，第二模块 12 释放机器虚拟机中用于描述旧的 vhost-net 模块的数据结构内存，这些数据结构及其他代码用于将 vhost-net 模块以设备的形式呈现给机器虚拟机使用。此后，第二模块 12 关闭旧的 vhost-net 模块的用户态接口导出设备的句柄（/dev/vhost-net），
20 以彻底移除旧的 vhost-net 模块。

可选地，第三模块 13 还用于：初始化第二半虚拟化网卡内核加速模块的资源。

在此，在插入新的 vhost-net 模块后，第三模块 13 打开新的 vhost-net 模块的用户态接口导出设备的句柄，并调用用户态下 QEMU 接口，在内存中初始化描述新的 vhost-net 模块的数据结构。

25 具体地，第四模块 14 包括：第四一子模块 141、第四二子模块 142、第四三子模块 143 和第四四子模块 144。

其中，第四一子模块 141，用于第二半虚拟化网卡内核加速模块创建内核线程；第四二子模块 142，用于第二半虚拟化网卡内核加速模块内核线程获取共享内存的使用状态；第四三子模块 143，用于第二半虚拟化网卡内核加速模块内核线程控制对共享内存
30 中网络数据包的发送和/或接收；第四四子模块 144，用于开启所述第二半虚拟化网卡内

核加速模块内核线程的轮询。

在此，第四一子模块 141 使初始化后的新的 vhost-net 模块创建内核线程；第四二子模块 142 使新的 vhost-net 模块创建的内核线程获取共享内存（vring）的使用状态；第四四子模块 144 开启新的 vhost-net 模块内核线程在 TAP 设备上的收/发网络数据包的轮询（poll）。第四三子模块 143 使新的 vhost-net 模块内核线程控制对共享内存中网络数据包的发送和/或接收。此后，获取共享内存使用状态的内核线程依照共享内存、TAP 设备、网桥及物理网卡的顺序对网络数据包进行接收和发送。

图 3 示出根据本申请一个方面的半虚拟化网卡内核加速模块升级方法的流程示意图。如图 3 所示，该半虚拟化网卡内核加速模块升级方法包括：

- 10 步骤 S301，将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输；
- 步骤 S302，移除第一半虚拟化网卡内核加速模块；
- 步骤 S303，插入第二半虚拟化网卡内核加速模块；
- 步骤 S304，将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输。

- 15 在此，半虚拟化网卡内核加速模块可以采用 KVM 虚拟化的 vhost-net 模块，vhost-net 模块用于机器虚拟机（QEMU）半虚拟化网卡（virtio-net）加速。第一半虚拟化网卡内核加速模块是指升级前旧的 vhost-net 模块，第二半虚拟化网卡内核加速模块是指升级后新的 vhost-net 模块。步骤 S301 将旧的 vhost-net 模块的网络流量切换至用户态进行传输。在用户态下，虚拟机在发送或接收数据包时，首先陷入内核态，然后再返回到用户态的
- 20 QEMU 进程进行处理，最后通过 TAP 设备经由网桥、物理网卡进行数据包的发送或接收。在将旧的 vhost-net 模块的网络流量切换至用户态进行转发后，可以移除旧的 vhost-net 模块而网络服务不被中断。此后插入新的 vhost-net 模块。将用户态进行转发的网络流量切换至内核态由升级后新的 vhost-net 模块进行传输。内核态下，虚拟机陷入内核态后，不需要再返回到用户态，而由新的 vhost-net 模块内核线程进行处理，通过 TAP 设备经由
- 25 网桥、物理网卡进行数据包的发送或接收。通过用户态和内核态的互相切换，从而实现了在移除旧的 vhost-net 模块并插入新的升级后的 vhost-net 模块时，不会中断网络服务。

具体地，步骤 S301 将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输包括：

- 步骤 S3011，机器虚拟机进程获取共享内存的使用状态；
- 30 步骤 S3012，所述机器虚拟机进程控制对共享内存中网络数据包的发送和/或接收。

具体地, 步骤 S301 将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输还包括:

步骤 S3013, 停止所述第一半虚拟化网卡内核加速模块内核线程的轮询。

在此, 停止旧的 vhost-net 模块内核线程在 TAP 设备上的收/发网络数据包的轮询 (poll), 开启 QEMU 进程对 TAP 设备上的可收/发网络数据包的轮询, 并且控制 QEMU 进程获取共享内存 (vring) 的使用状态, 由 QEMU 进程处理切换后的网络流量的收/发网络数据包, 将旧的 vhost-net 模块的网络流量切换至用户态进行处理, 为移除旧的 vhost-net 模块做好准备。其中, 共享内存用于缓存网络数据包。

具体地, 步骤 S302, 移除第一半虚拟化网卡内核加速模块还包括:

10 清理第一半虚拟化网卡内核加速模块的资源。

在此, 释放机器虚拟机中用于描述旧的 vhost-net 模块的数据结构内存, 这些数据结构及其他代码用于将 vhost-net 模块以设备的形式呈现给机器虚拟机使用。此后, 关闭旧的 vhost-net 模块的用户态接口导出设备的句柄(/dev/vhost-net), 以彻底移除旧的 vhost-net 模块。

15 具体地, 步骤 S303, 插入第二半虚拟化网卡内核加速模块还包括:

初始化第二半虚拟化网卡内核加速模块的资源。

在此, 在插入新的 vhost-net 模块后, 打开新的 vhost-net 模块的用户态接口导出设备的句柄, 并调用用户态下 QEMU 接口, 在内存中初始化描述新的 vhost-net 模块的数据结构。

20 具体地, 步骤 S304, 将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输包括:

步骤 S3041, 所述第二半虚拟化网卡内核加速模块创建内核线程;

步骤 S3042, 所述第二半虚拟化网卡内核加速模块内核线程获取共享内存的使用状态;

25 步骤 S3043, 所述第二半虚拟化网卡内核加速模块内核线程控制对共享内存中网络数据包的发送和/或接收。

具体地, 步骤 S304 将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输还包括:

步骤 S3044, 开启所述第二半虚拟化网卡内核加速模块内核线程的轮询。

30 在此, 新的 vhost-net 模块创建内核线程并控制该内核线程获取当前共享内存 (vring)

的使用状态，并且新的 vhost-net 模块开启内核线程在 TAP 设备上的收/发网络数据包的轮询（poll），以将用户态转发的网络流量切换至新的 vhost-net 模块。此后，获取共享内存使用状态的内核线程依照共享内存、TAP 设备、网桥及物理网卡的顺序对网络数据包进行接收和发送。

5 图 4 示出根据本申请又一个方面的半虚拟化网卡内核加速模块升级方法的流程示意图。如图 4 所示，该半虚拟化网卡内核加速模块升级方法包括：

步骤 S401，停止所述第一半虚拟化网卡内核加速模块内核线程的轮询。

在此，停止旧的 vhost-net 模块的内核线程在 TAP 设备上的收/发网络数据包的 poll（轮询）。

10 步骤 S402，机器虚拟机进程获取共享内存的使用状态。

在此，开启 QEMU 进程对 TAP 设备上的可收/发网络数据包的 poll（轮询），并且控制 QEMU 进程获取共享内存（vring）的使用状态。

步骤 S403，机器虚拟机进程控制对共享内存中网络数据包的发送和/或接收。

15 在此，由 QEMU 进程处理切换后的网络流量的收/发网络数据包，将旧的 vhost-net 模块的网络流量切换至用户态进行处理，为移除旧的 vhost-net 模块做好准备。

步骤 S404，移除第一半虚拟化网卡内核加速模块并清理第一半虚拟化网卡内核加速模块的资源。

20 在此，释放机器虚拟机中用于描述旧的 vhost-net 模块的数据结构内存，这些数据结构及其他代码用于将 vhost-net 模块以设备的形式呈现给机器虚拟机使用。此后，关闭旧的 vhost-net 模块的用户态接口导出设备的句柄(/dev/vhost-net)，以彻底移除旧的 vhost-net 模块。

步骤 S405，插入第二半虚拟化网卡内核加速模块并初始化第二半虚拟化网卡内核加速模块的资源。

25 在此，在插入新的 vhost-net 模块后，打开新的 vhost-net 模块的用户态接口导出设备的句柄，并调用用户态下 QEMU 接口，在内存中初始化描述新的 vhost-net 模块的数据结构。

步骤 S406，所述第二半虚拟化网卡内核加速模块创建内核线程。

步骤 S407，所述第二半虚拟化网卡内核加速模块内核线程获取共享内存的使用状态。

30 在此，新的 vhost-net 模块创建内核线程，内核线程获取共享内存使用状态。

步骤 S408, 开启所述第二半虚拟化网卡内核加速模块内核线程的轮询。

在此, 开启新的 vhost-net 模块内核线程在 TAP 设备上的收/发网络数据包的 poll (轮询)。

步骤 S409, 所述第二半虚拟化网卡内核加速模块内核线程控制对共享内存中网络数据包的发送和/或接收。

在此, 将用户态转发的网络流量切换至新的 vhost-net 模块, 依照共享内存、TAP 设备、网桥及物理网卡的顺序对网络数据包进行接收和发送。

需要注意的是, 本申请可在软件和/或软件与硬件的组合物中被实施, 例如, 可采用专用集成电路 (ASIC)、通用目的计算机或任何其他类似硬件设备来实现。在一个实施例中, 本申请的软件程序可以通过处理器执行以实现上文所述步骤或功能。同样地, 本申请的软件程序 (包括相关的数据结构) 可以被存储到计算机可读记录介质中, 例如, RAM 存储器, 磁或光驱动器或软磁盘及类似设备。另外, 本申请的一些步骤或功能可采用硬件来实现, 例如, 作为与处理器配合从而执行各个步骤或功能的电路。

另外, 本申请的一部分可被应用为计算机程序产品, 例如计算机程序指令, 当其被计算机执行时, 通过该计算机的操作, 可以调用或提供根据本申请的方法和/或技术方案。而调用本申请的方法的程序指令, 可能被存储在固定的或可移动的记录介质中, 和/或通过广播或其他信号承载媒体中的数据流而被传输, 和/或被存储在根据所述程序指令运行的计算机设备的工作存储器中。在此, 根据本申请的一个实施例包括一个装置, 该装置包括用于存储计算机程序指令的存储器和用于执行程序指令的处理器, 其中, 当该计算机程序指令被该处理器执行时, 触发该装置运行基于前述根据本申请的多个实施例的方法和/或技术方案。

对于本领域技术人员而言, 显然本申请不限于上述示范性实施例的细节, 而且在不背离本申请的精神或基本特征的情况下, 能够以其他的具体形式实现本申请。因此, 无论从哪一点来看, 均应将实施例看作是示范性的, 而且是非限制性的, 本申请的范围由所附权利要求而不是上述说明限定, 因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本申请内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。此外, 显然“包括”一词不排除其他单元或步骤, 单数不排除复数。装置权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第一, 第二等词语用来表示名称, 而并不表示任何特定的顺序。

权 利 要 求 书

1. 一种半虚拟化网卡内核加速模块升级方法，其中，所述方法包括：

将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输；

移除第一半虚拟化网卡内核加速模块；

5 插入第二半虚拟化网卡内核加速模块；

将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输。

2. 根据权利要求 1 所述的方法，其中，所述将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输包括：

机器虚拟机进程获取共享内存的使用状态；

10 所述机器虚拟机进程控制对共享内存中网络数据包的发送和/或接收。

3. 根据权利要求 2 所述的方法，其中，所述将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传输还包括：

停止所述第一半虚拟化网卡内核加速模块内核线程的轮询。

4. 根据权利要求 1 至 3 中任一项所述的方法，其中，所述移除第一半虚拟化网卡
15 内核加速模块包括：

清理所述第一半虚拟化网卡内核加速模块的资源。

5. 根据权利要求 1 至 4 中任一项所述的方法，其中，所述插入第二半虚拟化网卡
内核加速模块包括：

初始化所述第二半虚拟化网卡内核加速模块的资源。

20 6. 根据权利要求 1 至 5 中任一项所述的方法，其中，所述将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输包括：

所述第二半虚拟化网卡内核加速模块创建内核线程；

所述第二半虚拟化网卡内核加速模块内核线程获取共享内存的使用状态；

25 所述第二半虚拟化网卡内核加速模块内核线程控制对共享内存中网络数据包的发送和/或接收。

7. 根据权利要求 6 所述的方法，其中，所述将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模块进行传输还包括：

开启所述第二半虚拟化网卡内核加速模块内核线程的轮询。

8. 一种半虚拟化网卡内核加速模块升级装置，其中，所述装置包括：

30 第一模块，用于将第一半虚拟化网卡内核加速模块的网络流量切换至用户态进行传

输；

第二模块，用于移除第一半虚拟化网卡内核加速模块；

第三模块，用于插入第二半虚拟化网卡内核加速模块；

第四模块，用于将用户态进行传输的网络流量切换至第二半虚拟化网卡内核加速模

5 块进行传输。

9. 根据权利要求 8 所述的装置，其中，所述第一模块包括：

第一一子模块，用于机器虚拟机进程获取共享内存的使用状态；

第一二子模块，用于所述机器虚拟机进程控制对共享内存中网络数据包的发送和/或接收。

10 10. 根据权利要求 9 所述的装置，其中，所述第一模块还包括：

第一三子模块，用于停止所述第一半虚拟化网卡内核加速模块内核线程的轮询。

11. 根据权利要求 8 至 10 中任一项所述的装置，其中，所述第二模块还用于：清理所述第一半虚拟化网卡内核加速模块的资源。

12. 根据权利要求 8 至 11 中任一项所述的装置，其中，所述第三模块还用于：15 初始化所述第二半虚拟化网卡内核加速模块的资源。

13. 根据权利要求 8 至 12 中任一项所述的装置，其中，所述第四模块包括：

第四一子模块，用于第二半虚拟化网卡内核加速模块创建内核线程；

第四二子模块，用于第二半虚拟化网卡内核加速模块内核线程获取共享内存的使用状态；

20 第四三子模块，用于第二半虚拟化网卡内核加速模块内核线程控制对共享内存中网络数据包的发送和/或接收。

14. 根据权利要求 13 所述的装置，其中，所述第四模块包括：

第四四子模块，用于开启所述第二半虚拟化网卡内核加速模块内核线程的轮询。

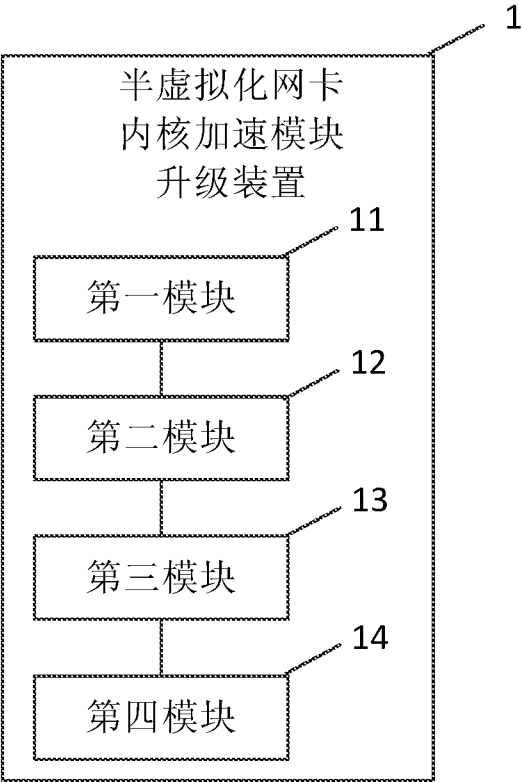


图 1

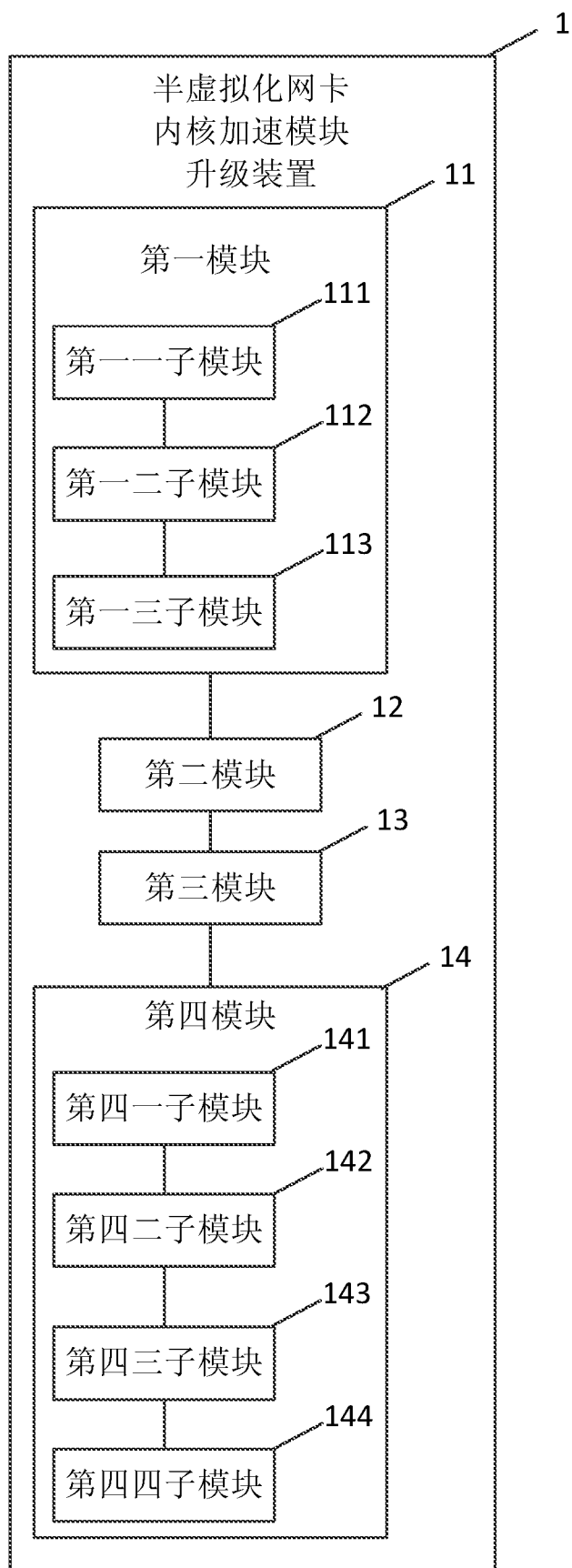


图 2

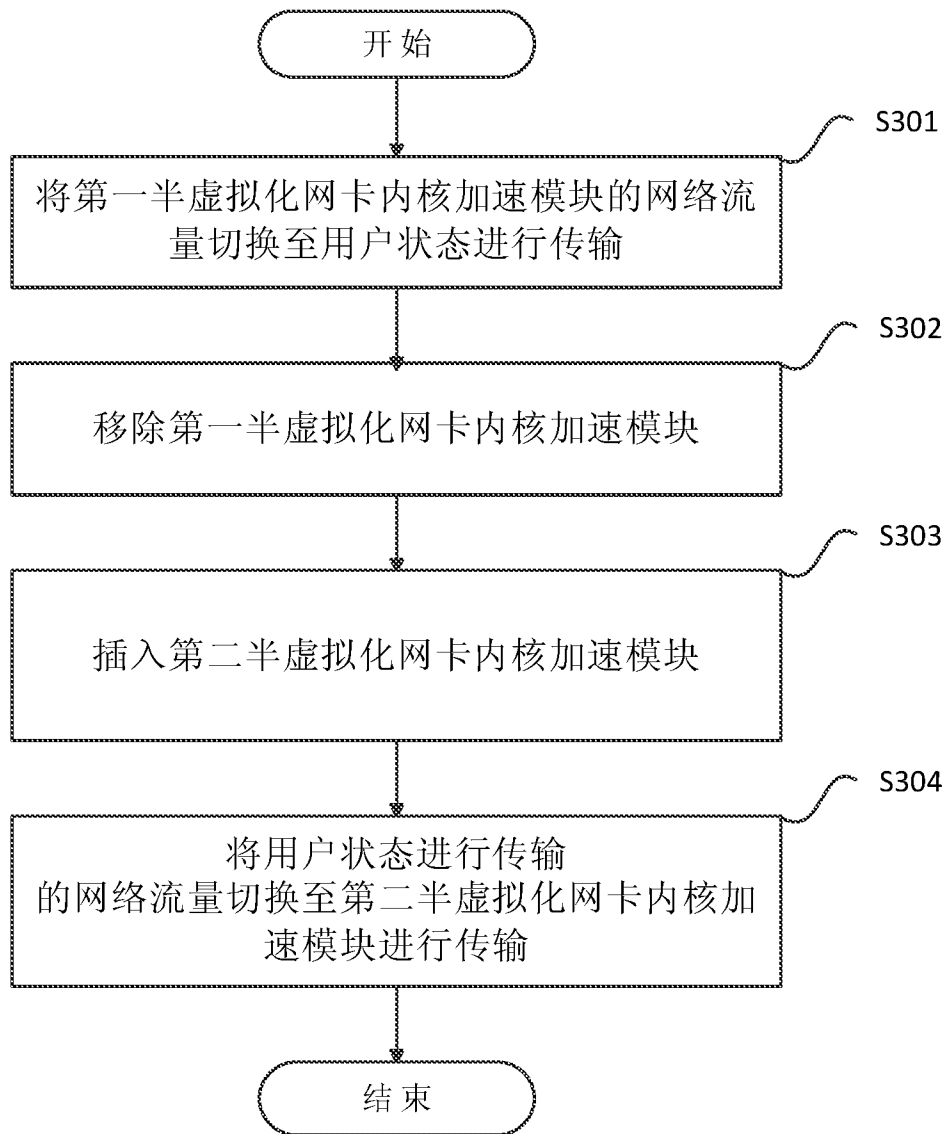


图 3

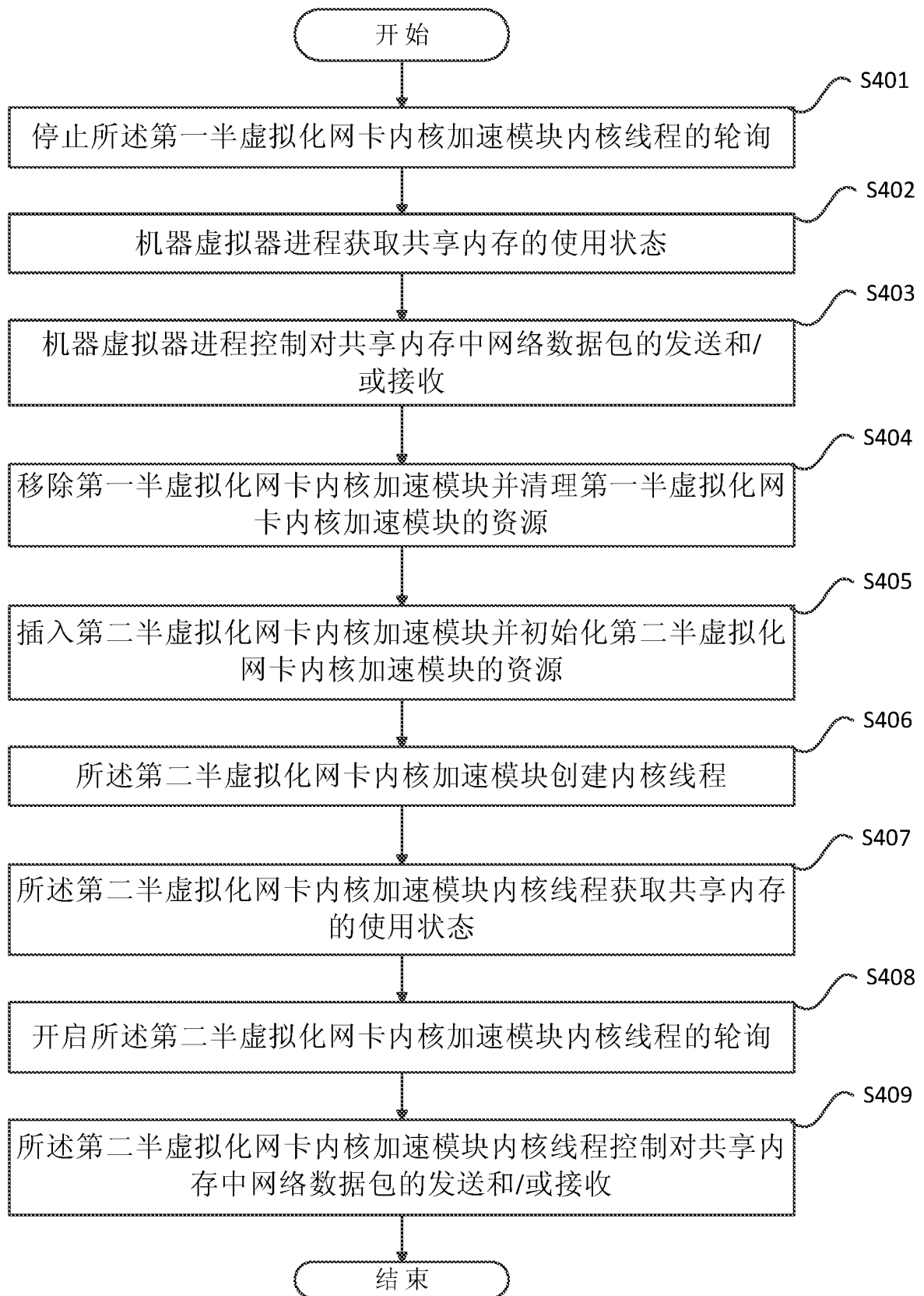


图 4

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/081953

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/445 (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)

IPC: G06F 9/-

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)

CNABS, CNTXT, SIPOABS, DWPI, CNKI: network, card, virtual, update, upgrade, switch, user mode, kernel, remove, insert, transmit, transfer

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 104423978 A (WUHAN HONGXINTONG TECHNOLOGY CO., LTD.) 18 March 2015 (18.03.2015) see the whole document	1-14
A	CN 101252464 A (HUAWEI TECHNOLOGIES CO., LTD.) 27 August 2008 (27.08.2008) see the whole document	1-14
A	WO 2009086681 A1 (ZTE CORP. et al.) 16 July 2009 (16.07.2009) see the whole document	1-14
A	CN 201332404 Y (SHENZHEN HUAWEI COMMUNICATION TECHNOLOGY CO., LTD.) 21 October 2009 (21.10.2009) see the whole document	1-14

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“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
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	“&” document member of the same patent family

Date of the actual completion of the international search 28 July 2016	Date of mailing of the international search report 15 August 2016
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer DU, Yi Telephone No. (86-10) 62411635

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/CN2016/081953

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104423978 A	18 March 2015	None	
CN 101252464 A	27 August 2008	CN 101252464 B	13 April 2011
WO 2009086681 A1	16 July 2009	None	
CN 201332404 Y	21 October 2009	None	

国际检索报告

国际申请号

PCT/CN2016/081953

A. 主题的分类

G06F 9/445 (2006.01) i

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

B. 检索领域

检索的最低限度文献 (标明分类系统和分类号)

G06F 9/-

包含在检索领域中的除最低限度文献以外的检索文献

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))

CNABS, CNTXT, SIPOABS, DWPI, CNKI: 网卡, 虚拟, 网络, 升级, 更新, 切换, 用户态, 内核, 移除, 插入, 传输, network, card, virtual, update, upgrade, switch, user mode, kernel, remove, insert, transmit, transfer

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 104423978 A (武汉鸿信通科技有限公司) 2015年 3月 18日 (2015 - 03 - 18) 见全文	1-14
A	CN 101252464 A (华为技术有限公司) 2008年 8月 27日 (2008 - 08 - 27) 见全文	1-14
A	WO 2009086681 A1 (中兴通讯股份有限公司 等) 2009年 7月 16日 (2009 - 07 - 16) 见全文	1-14
A	CN 201332404 Y (深圳华为通信技术有限公司) 2009年 10月 21日 (2009 - 10 - 21) 见全文	1-14

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“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)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“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

“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

“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

“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

“&” 同族专利的文件

国际检索实际完成的日期

2016年 7月 28日

国际检索报告邮寄日期

2016年 8月 15日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局 (ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

授权官员

杜轶

传真号 (86-10) 62019451

电话号码 (86-10) 62411635

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2016/081953

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	104423978	A	2015年 3月 18日	无			
CN	101252464	A	2008年 8月 27日	CN	101252464	B	2011年 4月 13日
WO	2009086681	A1	2009年 7月 16日	无			
CN	201332404	Y	2009年 10月 21日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)