(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
2018 年 2 月 8 日 (08.02.2018)

WIPO | PCT

(10) 国际公布号
WO 2018/023499 A1

(51) 国际专利分类号:
G06F 13/14 (2006.01) *H04L 12/931* (2013.01)

(21) 国际申请号: PCT/CN2016/093098

(22) 国际申请日: 2016 年 8 月 3 日 (03.08.2016)

(25) 申请语言: 中文

(26) 公布语言: 中文

(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: 吴 天 议 **(WU, Tianyi)**; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 张 忠 军 **(ZHANG, Zhongjun)**; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 甘 涛 **(GAN, Tao)**; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (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。

(54) Title: NETWORK INTERFACE CARD, COMPUTER DEVICE AND DATA PACKET PROCESSING METHOD

(54) 发明名称: 网络接口卡、计算设备以及数据包处理方法

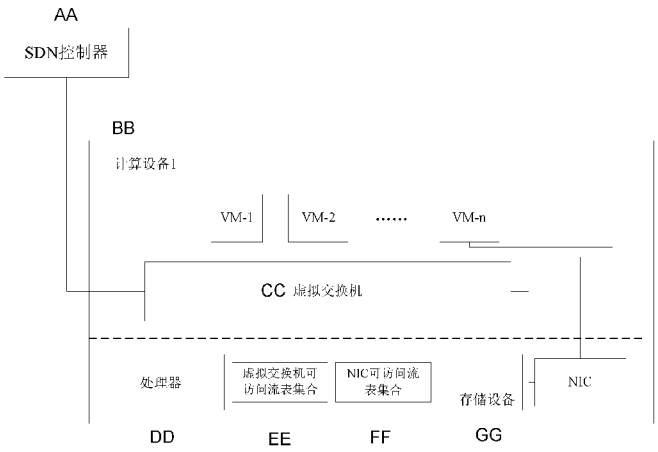

图 2c

AA SDN controller
BB Computer device 1
CC Virtual switch
DD Processor
EE Virtual switch accessible flow table set
FF NIC accessible flow table set
GG Storage device

(57) Abstract: Disclosed is a data packet processing method, used in a computer device in a software-defined network (SDN). After receiving a data packet of a data stream, a network interface card (NIC) in the computer device queries a flow table set according to matching information of the data packet, and if a flow table is matched in the flow table set, processes the data packet according to the flow table, and if no flow table can be matched in the flow table set, sends the data packet to a virtual switch, so that the virtual switch obtains a flow table corresponding to the data stream in which the data packet is located, and stores the flow table in the flow table set

WO 2018/023499 A1

[见续页]

(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)。

根据细则4.17的声明：

— 关于申请人有权申请并被授予专利(细则4.17(ii))

本国际公布：

— 包括国际检索报告(条约第21条(3))。

so that the NIC can directly process subsequent data packets of the data stream in which the data packet is located. The method reduces the work load of the virtual switch, and increases the work efficiency of the computer device.

(57) 摘要：本申请公开了一种数据包处理方法，该方法运用于软件定义网络SDN中的计算设备。计算设备中的NIC接收到数据流的数据包后，根据数据包的匹配信息查询流表集合，如果在流表集合中匹配上流表，则根据该流表处理该数据包，如果在流表集合中无法匹配任一流表，则NIC将该数据包发送至虚拟交换机，以使虚拟交换机获取该数据包所在数据流对应的流表，并将该流表存入流表集合以供该NIC能够直接处理该数据包所在数据流的后续数据包。本申请提供的方法降低了虚拟交换机的工作负担，提升了计算设备的工作效率。

网络接口卡、计算设备以及数据包处理方法

技术领域

本申请涉及计算机技术领域，尤其涉及一种网络接口卡（英文全称：network interface card，缩写：NIC），一种用于处理数据包的计算设备，以及该NIC、该计算设备分别执行的数据包处理方法。

背景技术

云计算环境中，由于需要对数量较高的用户提供服务，因此用于提供云服务的数据中心中的计算设备的数量往往较多，而每个计算设备上又运行了多个虚拟机（英文全称：virtual machine，缩写：VM），如图1中的VM-1至VM-n。VM与其他计算设备上运行的VM或同一计算设备上的VM之间通过虚拟交换机（英文全称：virtual switch，缩写：VS）通信，软件定义网络（英文全称：software defined networking，缩写：SDN）控制器对各个计算设备上的虚拟交换机集中进行控制。当前常见的虚拟交换机包括open vSwitch，SDN控制器通常通过OpenFlow协议定义的流表（英文全称：flow table）对各个虚拟交换机进行控制。

每个计算设备上的硬件资源至少需要支持运行多个VM、虚拟交换机以及虚拟机监视器（英文全称：virtual machine monitor），虚拟机监视器又称为虚拟机管理器（英文全称：virtual machine manager）或管理程序（英文全称：hypervisor）。每台计算设备的硬件资源有限，如果负担了数据交换任务的虚拟交换机占用的硬件资源太多，则容易影响计算设备上VM的运行，降低工作效率。

发明内容

本申请提供了一种数据包处理方法，以提升数据包处理效率。

本申请的第一方面，提供了一种数据包处理方法，该方法应用于计算设备，该计算设备包括网络接口卡NIC和主机，该NIC与该主机建立通信连接，该主机运行虚拟机VM，该方法包括：该NIC接收数据流的第一数据包；该NIC根据该

第一数据包的匹配信息查询流表集合；在无法匹配到该数据流对应的流表情况下，该NIC向该主机上运行的虚拟交换机转发该第一数据包；其中，该虚拟交换机在接收到该第一数据包后，从SDN控制器获取该数据流对应的流表，以便于该数据流对应的流表被加入到该流表集合。

该虚拟交换机通过该第一数据包获取了该数据流对应的流表后，由该虚拟交换机或主机上运行的监控模块将该数据流对应的流表存入该流表集合。

实际运行中，该 NIC 接收一个数据包后，根据该数据包的匹配信息查询流表集合，如果无法获取该数据包所在数据流对应的流表，说明该数据包为该数据包所在数据流的首个数据包，或该数据包不是所在数据流的首个数据包，但该流表集合中该数据流对应的流表已经老化。

该数据包处理方法，由该 NIC 执行流表与数据包的匹配动作，提升了数据包的处理效率，并且将无法匹配到流表的数据包发送至虚拟交换机，以获取对应的流表用于 NIC 对该数据流后续的数据包的处理。

结合第一方面，在第一方面的第一种实现方式中，该主机上运行的虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应该主机上运行的一个VM，该NIC接收该第一数据包前，该方法还包括：该NIC接收该虚拟交换机端口的配置信息；该NIC根据该虚拟交换机端口的配置信息，在该NIC上配置至少一个NIC端口，每个NIC端口通过单根-输入/输出虚拟化（英文全称：single root-input/output virtualization，缩写：SR-I/OV）技术与该主机上运行的一个VM连接。

虚拟交换机端口以及NIC端口的配置可以在第一方面提供的数据包处理方法之前完成，并且可以由该主机上运行的NIC驱动向该NIC发送该虚拟交换机端口的配置信息。

结合第一方面的第一种实现方式，在第一方面的第二种实现方式中，该数据流对应的流表包括该数据流的数据包的路由信息，在向该虚拟交换机转发该第一数据包之后，该方法还包括：该NIC，根据该第一数据包的匹配信息查询该流表集合，获取该数据流对应的流表，并根据该数据流的数据包的路由信息转发该第一数据包至目的VM。该数据流的数据包的路由信息指示该目的VM对应

的NIC端口。

该虚拟交换机或主机上运行的监控模块将该数据流对应的流表存入该流表集合后，由该NIC根据该第一数据包的匹配信息查询该流表集合，此时该流表集合中已经存有该数据流对应的流表。该NIC可以在将该第一数据包发送至该虚拟交换机之后，定期将该第一数据包与该流表集合中的流表进行匹配，或者该数据流对应的流表存入该流表集合后，向该NIC发送通知消息，指示该NIC执行该第一数据包与该流表集合中的流表的匹配。

该实现方式无须该虚拟交换机执行该第一数据包与流表的匹配，降低了该虚拟交换机的工作负担。

结合第一方面的第一种实现方式，在第一方面的第三种实现方式中，该数据流对应的流表包括该数据流的数据包的路由信息，在向该虚拟交换机转发该第一数据包之后，该方法还包括：该NIC接收该虚拟交换机返回的第一数据包，该返回的第一数据包中包含该对应于目的VM的端口标识，该端口标识为虚拟交换机端口标识或NIC端口标识；该NIC根据该端口标识，转发该第一数据包到该目的VM，该端口标识由该虚拟交换机根据该数据流的数据包的路由信息添加。

该实现方式无须该NIC与该第一数据包的匹配，提升了数据包的处理效率。

结合第一方面的第一种实现方式，在第一方面的第四种实现方式中，该虚拟交换机与该NIC通过至少一个队列通信，每个队列对应该主机上运行的一个VM；在向该虚拟交换机转发该第一数据包之后，该方法还包括：该NIC从该第一数据包的目的VM对应的队列接收该第一数据包；该NIC根据该目的VM对应的队列的队列信息，从该目的VM对应的NIC端口转发该第一数据包至该目的VM。

该虚拟交换机获取该数据流的数据包的路由信息后，根据该数据流的数据包的路由信息将该第一数据包存入该目的VM对应的队列中。该NIC从该目的VM对应的队列接收该第一数据包，主机上运行的NIC驱动向该NIC发送该目的VM对应的队列的队列信息，该队列信息用于通知该NIC该第一数据包从该目的VM对应的队列获得，该NIC根据预存的队列与NIC端口的对应关系，从该目的VM对应的NIC端口转发该第一数据包至该目的VM。

该实现方式无须该NIC与该第一数据包的匹配，与前述第三种实现方式相比，

也无须该NIC对端口标识的转换，进一步提升了数据包的处理效率。

本申请的第二方面，提供了一种NIC，该NIC用于执行本申请第一方面提供的数据包处理方法。该NIC包括：主机接口、网络接口和处理芯片，该网络接口用于与外部网络通信且该网络接口与该处理芯片建立通信连接，该主机接口用于与主机通信且该主机接口与该处理芯片建立通信连接，该主机运行VM；该网络接口，用于接收数据流的第一数据包；该处理芯片，用于根据该第一数据包的匹配信息查询流表集合，在无法匹配到该数据流对应的流表情况下，通过该主机接口向该主机上运行的虚拟交换机转发该第一数据包；其中，该虚拟交换机在接收到该第一数据包后，从SDN控制器获取该数据流对应的流表，以便于该数据流对应的流表被加入到该流表集合。

该 NIC 能够执行流表与数据包的匹配动作，提升了数据包的处理效率，并且该 NIC 将无法匹配到流表的数据包发送至虚拟交换机，以获取对应的流表用于 NIC 对该数据流后续的数据包的处理。

结合第二方面，在第二方面的第一种实现方式中，该主机上运行的虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应该主机上运行的一个VM；该处理芯片，还用于接收该虚拟交换机端口的配置信息，根据该虚拟交换机端口的配置信息，在该NIC上配置至少一个NIC端口，每个NIC端口通过SR-I/OV技术与该主机上运行的一个VM连接。

结合第二方面的第一种实现方式，在第二方面的第二种实现方式中，该处理芯片，还用于根据该第一数据包的匹配信息查询该流表集合，获取该数据流对应的流表，该数据流对应的流表包括该数据流的数据包的路由信息，并根据该数据流的数据包的路由信息转发该第一数据包至目的VM。

该实现方式中，无须该虚拟交换机与该第一数据包的匹配，降低了该虚拟交换机的工作负担。

结合第二方面的第一种实现方式，在第二方面的第三种实现方式中，该处理芯片，还用于接收该虚拟交换机返回的第一数据包，该返回的第一数据包中包含对应于目的VM的端口标识，该端口标识为虚拟交换机端口标识或NIC端口

标识，该端口标识由该虚拟交换机根据该数据流的数据包的路由信息添加，该数据流对应的流表包括该数据流的数据包的路由信息；还用于根据该端口标识，转发该第一数据包到该目的VM。

该实现方式中，该NIC无须将该第一数据包与流表进行匹配，提升了数据包的处理效率。

结合第二方面的第一种实现方式，在第二方面的第四种实现方式中，该虚拟交换机与该NIC通过至少一个队列通信，每个队列对应该主机上运行的一个VM；该处理芯片，还用于从该第一数据包的目的VM对应的队列接收该第二数据包，并根据该目的VM对应的队列的队列信息，从该目的VM对应的NIC端口转发该第二数据包至该目的VM。

该实现方式中，该NIC无须将该第一数据包与流表进行匹配，并且与前述第三种实现方式相比，该NIC也无须对该端口标识进行转换，进一步提升了数据包的处理效率。

本申请第三方面提供了一种数据包处理方法，该方法应用于计算设备，该计算设备包括网络接口卡NIC和主机，该NIC与该主机建立通信连接，该主机运行虚拟机VM，该方法包括：该NIC接收数据流的第二数据包；该NIC根据该第二数据包的匹配信息查询流表集合，获取该数据流对应的流表，该数据流对应的流表包括该数据流的数据包的路由信息；该NIC根据该数据流的数据包的路由信息，转发该第二数据包到目的VM。

该数据包处理方法，由该 NIC 执行流表与数据包的匹配动作，提升了数据包的处理效率。

结合第三方面，在第三方面的第一种实现方式中，该NIC通过SR-I/OV技术与该主机上运行的VM连接；该NIC根据该数据流的数据包的路由信息，通过该NIC与该目的VM的连接转发该第二数据包到该目的VM。

结合第三方面的第一种实现方式，在第三方面的第二种实现方式中，该主机上运行的虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应该主机上运行的一个VM；该NIC接收数据流的第二数据包前，该方法还包

括：该NIC接收该虚拟交换机端口的配置信息；根据该虚拟交换机端口的配置信息，在该NIC上配置至少一个NIC端口，每个NIC端口通过SR-I/OV技术与该主机上运行的一个VM连接，该数据流的数据包的路由信息指示该目的VM对应的NIC端口。

结合第三方面的第二种实现方式，在第三方面的第三种实现方式中，该NIC接收数据流的第二数据包前，该方法还包括：该NIC接收该数据流的第三数据包；该NIC根据该第三数据包的匹配信息查询该流表集合；在无法匹配到该数据流对应的流表情况下，该NIC向该主机上运行的虚拟交换机转发该第三数据包；其中，该虚拟交换机在接收到该第三数据包后，从SDN控制器获取该数据流对应的流表，以便于该数据流对应的流表被加入到该流表集合。

结合第三方面的第三种实现方式，在第三方面的第四种实现方式中，在向该虚拟交换机转发该第三数据包之后，该方法还包括：该NIC根据该第三数据包的匹配信息查询该流表集合，获取该数据流对应的流表，并根据该数据流的数据包的路由信息转发该第三数据包至该目的VM。

结合第三方面的第三种实现方式，在第三方面的第五种实现方式中，在向该虚拟交换机转发该第三数据包之后，该方法还包括：该NIC接收该虚拟交换机返回的第三数据包，该返回的第三数据包中包含对应于该目的VM的端口标识，该端口标识为虚拟交换机端口标识或NIC端口标识，该端口标识由该虚拟交换机根据该数据流的数据包的路由信息添加；该NIC根据该端口标识，转发该第三数据包到该目的VM。

结合第三方面的第三种实现方式，在第三方面的第六种实现方式中，该虚拟交换机与该NIC通过至少一个队列通信，每个队列对应该主机上运行的一个VM；在向该虚拟交换机转发该第三数据包之后，该方法还包括：该NIC从该目的VM对应的队列接收该第三数据包；该NIC根据该目的VM对应的队列的队列信息，从该目的VM对应的NIC端口转发该第三数据包至该目的VM。

结合第三方面或第三方面的前述任一种实现方式，在第三方面的第七种实现方式中，该NIC接收数据流的第二数据包包括：该NIC接收overlay类型数据包，该overlay类型数据包包括overlay头和该第二数据包，该overlay头包括虚拟可扩展

局域网VXLAN头，或使用通用路由的网络虚拟化NVGRE头，或无状态传输隧道STT头；该NIC剥去该overlay类型数据包的overlay头，获取该第二数据包。

需要说明的是，如果该数据流的数据包均采用了overlay技术，该NIC也需要对将该第三数据包对应的overlay头剥离后，才能获取该第三数据包。本实现方式中，示例性的提出了该NIC对该overlay类型数据包的剥离，实际上该NIC接收该数据流的其他overlay类型数据包后，也会将其overlay头剥离以获取内层的数据包。

该实现方式中，该NIC实现了对overlay头的剥离，降低了主机的工作负担。

结合第三方面或第三方面的前述任一种实现方式，在第三方面的第八种实现方式中，该NIC转发该第二数据包到目的VM之前，该方法还包括：该NIC对该第二数据包进行安全组检查，在确定该第二数据包的安全组检查通过之后，执行转发该第二数据包到该目的VM的步骤。

需要说明的是，如果该数据流的数据包均设置了安全组，，该NIC也需要对确定该第三数据包的安全组检查通过后，才转发该第三数据包。本实现方式中，示例性的提出了该NIC对该第二数据包的安全组检查，实际上该NIC转发该数据流的其他数据包之前，也需要对其进行安全组检查。

该实现方式中，该NIC还实现了对该第二数据包的安全组检查，提升了数据包收发的安全的同时，进一步降低了主机的工作负担。

本申请的第四方面提供了一种NIC，该NIC用于执行前述第三方面提供的数据包处理方法。该NIC包括：主机接口、网络接口和处理芯片，该网络接口用于与外部网络通信且该网络接口与该处理芯片建立通信连接，该主机接口用于与主机通信且该主机接口与该处理芯片建立通信连接，该主机运行VM；该网络接口，用于接收数据流的第二数据包；该处理芯片，用于根据该第二数据包的匹配信息查询流表集合，获取该数据流对应的流表，该数据流对应的流表包括该数据流的数据包的路由信息；以及根据该数据流的数据包的路由信息，转发该第二数据包到目的VM。

结合第四方面，在第四方面的第一种实现方式中，该NIC通过SR-I/OV技术与该主机上运行的VM连接；

该处理芯片，用于根据该数据流的数据包的路由信息，通过该NIC与该目的VM的连接转发该第二数据包到该目的VM。

结合第四方面的第一种实现方式，在第四方面的第二种实现方式中，该主机上运行的虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应该主机上运行的一个VM； 该处理芯片，还用于接收该虚拟交换机端口的配置信息，根据该虚拟交换机端口的配置信息在该NIC上配置至少一个NIC端口，每个NIC端口通过SR-I/OV技术与该主机上运行的一个VM连接，该数据流的数据包的路由信息指示该目的VM对应的NIC端口。

结合第四方面的第二种实现方式，在第四方面的第三种实现方式中，该网络接口，还用于接收该数据流的第三数据包；该处理芯片，还用于根据该第三数据包的匹配信息查询该流表集合，在无法匹配到该数据流对应的流表情况下，向该主机上运行的虚拟交换机转发该第三数据包；其中，该虚拟交换机在接收到该第三数据包后，从SDN控制器获取该数据流对应的流表，以便于该数据流对应的流表被加入到该流表集合。

结合第四方面的第三种实现方式，在第四方面的第四种实现方式中，该处理芯片，还用于根据该第三数据包的匹配信息查询该流表集合，获取该数据流对应的流表，并根据该数据流的数据包的路由信息转发该第三数据包至该目的VM。

结合第四方面的第三种实现方式，在第四方面的第五种实现方式中，该处理芯片，还用于接收该虚拟交换机返回的第三数据包，该返回的第三数据包中包含对应于该目的VM的端口标识，该端口标识为虚拟交换机端口标识或NIC端口标识，该端口标识由该虚拟交换机根据该数据流的数据包的路由信息添加；还用于根据该端口标识，转发该第三数据包到该目的VM。

结合第四方面的第三种实现方式，在第四方面的第六种实现方式中，该虚拟交换机与该NIC通过至少一个队列通信，每个队列对应该主机上运行的一个VM；该处理芯片，还用于从该目的VM对应的队列接收该第三数据包，并根据该目的VM对应的队列的队列信息，从该目的VM对应的NIC端口转发该第三数据包至该目的VM。

结合第四方面或第四方面的前述任一种实现方式，在第四方面的第七种实现方式中，该网络接口，用于接收overlay类型数据包，该overlay类型数据包包括overlay头和该第二数据包，该overlay头包括虚拟可扩展局域网VXLAN头，或使用通用路由的网络虚拟化NVGRE头，或无状态传输隧道STT头；该处理芯片，用于剥去该overlay类型数据包的overlay头，获取该第二数据包。

结合第四方面或第四方面的前述任一种实现方式，在第四方面的第八种实现方式中，该处理芯片转发该第二数据包到该目的VM之前，还用于对该第二数据包进行安全组检查，在确定该第二数据包的安全组检查通过之后，执行转发该第二数据包到该目的VM的步骤。

本申请的第五方面提供了一种计算设备，该计算设备包括网络接口卡NIC和主机，该NIC与该主机建立通信连接，该主机上运行虚拟机VM和虚拟交换机，该虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应该主机上运行的一个VM；该主机，用于向该NIC发送该虚拟交换机端口的配置信息；该NIC，用于根据该虚拟交换机端口的配置信息，在该NIC上配置至少一个NIC端口，每个NIC端口通过SR-I/OV技术与该主机上运行的一个VM连接；该NIC，还用于接收数据流的第二数据包，根据该第二数据包的匹配信息查询流表集合，获取该数据流对应的流表，该数据流对应的流表包括该数据流的数据包的路由信息，该数据流的数据包的路由信息指示目的VM对应的NIC端口，以及根据该数据流的数据包的路由信息，转发该第一数据包到该目的VM。

结合第五方面，在第五方面的第一种实现方式中，该NIC，还用于接收该数据流的第三数据包，根据该第三数据包的匹配信息查询该流表集合，在无法匹配到该数据流对应的流表情况下，向该主机转发该第三数据包；该主机，用于在接收到该第三数据包后，从SDN控制器获取该数据流对应的流表，以便于该数据流对应的流表被加入到该流表集合。

具体的，该NIC向该主机上运行的虚拟交换机发送该第三数据包。

结合第五方面的第一种实现方式，在第五方面的第二种实现方式中，该NIC，还用于根据该第三数据包的匹配信息查询该流表集合，获取该数据流对应的流

表，以及根据该数据流的数据包的路由信息转发该第三数据包至该目的VM。

该NIC在该数据流对应的流表被加入到该流表集合后，根据该第三数据包的匹配信息查询该流表集合。

结合第五方面的第一种实现方式，在第五方面的第三种实现方式中，该主机，还用于生成返回的第三数据包，该返回的第三数据包中包含对应于该目的VM的端口标识，该端口标识为虚拟交换机端口标识或NIC端口标识，该端口标识由该主机根据该数据流的数据包的路由信息添加；该NIC，还用于接收该返回的第三数据包，并根据该端口标识，转发该第三数据包到该目的VM。

结合第五方面的第一种实现方式，在第五方面的第四种实现方式中，该虚拟交换机与该NIC通过至少一个队列通信，每个队列对应该主机上运行的一个VM；该主机，还用于将该第三数据包发送至该目的VM对应的队列；该NIC，还用于从该目的VM对应的队列接收该第三数据包，以及根据该目的VM对应的队列的队列信息，从该目的VM对应的NIC端口转发该第三数据包至该目的VM。

该主机在接收到该NIC在发来的该第三数据包之后，将该第三数据包发送至该目的VM对应的队列。

结合第五方面或第五方面的前述任一种实现方式，在第五方面的第五种实现方式中，该NIC，具体用于接收overlay类型数据包，该overlay类型数据包包括overlay头和该第二数据包，该overlay头包括虚拟可扩展局域网VXLAN头，或使用通用路由的网络虚拟化NVGRE头，或无状态传输隧道STT头，以及剥去该overlay类型数据包的overlay头，获取该第二数据包。

结合第五方面或第五方面的前述任一种实现方式，在第五方面的第六种实现方式中，该NIC转发该第二数据包到目的VM之前，还用于对该第二数据包进行安全组检查，在确定该第二数据包的安全组检查通过之后，执行转发该第二数据包到该目的VM的步骤。

本申请的第六方面提供了一种数据包处理方法，该方法运用于前述第五方面提供的计算设备。该方法包括：主机向NIC发送该虚拟交换机端口的配置信息；该NIC根据该虚拟交换机端口的配置信息，在该NIC上配置至少一个NIC端口，

每个NIC端口通过SR-I/OV技术与该主机上运行的一个VM连接；该NIC接收数据流的第二数据包，根据该第二数据包的匹配信息查询流表集合，获取该数据流对应的流表，该数据流对应的流表包括该数据流的数据包的路由信息，该数据流的数据包的路由信息指示目的VM对应的NIC端口，以及根据该数据流的数据包的路由信息，转发该第二数据包到该目的VM。

结合第六方面，在第六方面的第一种实现方式中，该方法还包括：该NIC接收该数据流的第三数据包，根据该第三数据包的匹配信息查询该流表集合，在无法匹配到该数据流对应的流表情况下，向该主机转发该第三数据包；该主机在接收到该第三数据包后，从SDN控制器获取该数据流对应的流表，以便于该数据流对应的流表被加入到该流表集合。

结合第六方面的第一种实现方式，在第六方面的第二种实现方式中，在该数据流对应的流表被加入到该流表集合之后，该方法还包括：该NIC根据该第三数据包的匹配信息查询该流表集合，获取该数据流对应的流表，以及根据该数据流的数据包的路由信息转发该第三数据包至该目的VM。

结合第六方面的第一种实现方式，在第六方面的第三种实现方式中，在该NIC向该主机转发该第三数据包之后，该方法还包括：该主机生成返回的第三数据包，该返回的第三数据包中包含对应于该目的VM的端口标识，该端口标识为虚拟交换机端口标识或NIC端口标识，该端口标识由该主机根据该数据流的数据包的路由信息添加；该NIC接收该返回的第三数据包，并根据该端口标识，转发该第三数据包到该目的VM。

结合第六方面的第一种实现方式，在第六方面的第四种实现方式中，该虚拟交换机与该NIC通过至少一个队列通信，每个队列对应该主机上运行的一个VM；在该NIC向该主机转发该第三数据包之后，该方法还包括：该主机将该第三数据包发送至该目的VM对应的队列；该NIC从该目的VM对应的队列接收该第三数据包，以及根据该目的VM对应的队列的队列信息，从该目的VM对应的NIC端口转发该第三数据包至该目的VM。

结合第六方面或第六方面的前述任一种实现方式，在第六方面的第五种实现方式中，该NIC接收该第二数据包具体包括：该NIC接收overlay类型数据包，

该overlay类型数据包包括overlay头和该第二数据包，该overlay头包括虚拟可扩展局域网VXLAN头，或使用通用路由的网络虚拟化NVGRE头，或无状态传输隧道STT头，以及剥去该overlay类型数据包的overlay头，获取该第二数据包。

结合第六方面或第六方面的前述任一种实现方式，在第六方面的第六种实现方式中，该NIC转发该第二数据包到目的VM之前，该方法包括：该NIC对该第二数据包进行安全组检查，在确定该第二数据包的安全组检查通过之后，执行转发该第二数据包到该目的VM的步骤。

本申请的第七方面，提供了一种配置方法，该配置方法应用于主机，该主机与NIC建立通信连接，该主机运行VM、虚拟交换机和NIC驱动，该主机上运行的虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应该主机上运行的一个VM，该方法包括：该NIC驱动向该NIC发送该虚拟交换机端口的配置信息，该虚拟交换机端口的配置信息指示在该NIC上配置至少一个NIC端口，每个NIC端口通过SR-I/OV技术与该主机上运行的一个VM连接。

本申请的第八方面，提供了一种数据包处理方法，该方法应用于执行了本申请第七方面提供的配置方法的主机，该主机运行时执行本申请第五方面提供的数据包处理方法中主机侧的部分。该方法包括：该主机上运行的虚拟交换机接收数据流的第三数据包；该虚拟交换机从SDN控制器获取该数据流对应的流表；该虚拟交换机或该主机运行的监控模块将该数据流对应的流表加入到流表集合。

结合第八方面，在第八方面的第一种实现方式中，在该数据流对应的流表被加入到该流表集合后，该方法还包括：该虚拟交换机向该NIC发送通知消息，该通知消息用于通知该NIC该数据流对应的流表已经加入到流表集合，以使得该NIC根据该流表集合中的该数据流对应的流表，处理该第三数据包。

结合第八方面，在第八方面的第二种实现方式中，该数据流对应的流表包括该数据流的数据包的路由信息，该方法还包括：该虚拟交换机根据该数据流的数据包的路由信息生成返回的第三数据包，将该返回的第三数据包发送至该

NIC，该返回的第三数据包中包含对应于该目的VM的端口标识，该端口标识为虚拟交换机端口标识或NIC端口标识，以使该NIC根据该端口标识转发该第三数据包到该目的VM。

结合第八方面，在第八方面的第三种实现方式中，该虚拟交换机与该NIC通过至少一个队列通信，每个队列对应该主机上运行的一个VM；该方法还包括：该虚拟交换机将该第三数据包发送至该目的VM对应的队列；该NIC驱动将该目的VM对应的队列的队列信息发送至该NIC，以使该NIC根据该目的VM对应的队列的队列信息，从该目的VM对应的NIC端口转发该第三数据包至该目的VM。

本申请第九方面提供了一种主机，该主机包括处理器、存储器、总线，该处理器和该存储器通过该总线建立通信连接，该处理器运行时，执行前述第七方面提供的配置方法。

本申请第十方面提供了一种主机，该主机包括处理器、存储器、总线，该处理器和该存储器通过该总线建立通信连接，该处理器运行时，执行前述第八方面或第八方面的任一种实现方式提供的数据包处理方法。

本申请的第十一方面，提供了一种存储介质，该存储介质中存储了程序代码，该程序代码被计算设备运行时，执行第七方面提供的配置方法。该存储介质包括但不限于快闪存储器、硬盘（英文：hard disk drive，缩写：HDD）或固态硬盘（英文：solid state drive，缩写：SSD）。

本申请的第十二方面，提供了一种存储介质，该存储介质中存储了程序代码，该程序代码被计算设备运行时，执行第八方面或第八方面的任意一种实现方式提供的数据包处理方法。该存储介质包括但不限于快闪存储器、HDD或SSD。

本申请的第十三方面，提供了一种计算机程序产品，该计算机程序产品可以为一个软件安装包，该软件安装包被计算设备运行时，执行第七方面提供的

配置方法。

本申请的第十四方面，提供了一种计算机程序产品，该计算机程序产品可以为一个软件安装包，该软件安装包被计算设备运行时，执行第八方面或第八方面的任意一种实现方式提供的数据包处理方法。

附图说明

为了更清楚地说明本申请实施例的技术方案，下面将对实施例中所需要使用的附图作以简单地介绍，显而易见的，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。

图 1 为现有技术中数据中心架构的示意图；

图 2a 为本申请实施例提供的 SDN 架构的示意图；

图 2b 为现有技术中 SDN 中的计算设备的组织结构示意图；

图 2c 为本申请实施例提供的计算设备的组织结构示意图；

图 3 为本申请实施例提供的又一计算设备的组织结构示意图；

图 4 为本申请实施例提供的数据包处理方法的流程示意图；

图 5a 为本申请实施例提供的另一计算设备的组织结构示意图；

图 5b 为本申请实施例提供的另一计算设备的组织结构示意图；

图 5c 为本申请实施例提供的另一计算设备的组织结构示意图；

图 5d 为本申请实施例提供的另一计算设备的组织结构示意图；

图 6a 为本申请实施例提供的 NIC 的组织结构示意图；

图 6b 为本申请实施例提供的又一 NIC 的组织结构示意图；

图 6c 为本申请实施例提供的又一 NIC 的组织结构示意图。

具体实施方式

下面结合本申请实施例中的附图，对本申请实施例中的技术方案进行描述。

本申请中采用术语第一、第二、第三等来区分各个对象，例如第一数据包、第二数据包、第三数据包等，但各个“第一”、“第二”、“第三”之间不具有逻辑或时序上的依赖关系。

贯穿本说明书，数据包由匹配信息和载荷（英文全称：payload）构成。其中，匹配信息用于与流表的匹配域进行匹配。

贯穿本说明书，流表（英文全称：flow table）用于在 SDN 中控制数据流，也可以称为 SDN 流表，具体可以采用符合 OpenFlow 协议的流表或符合其他协议的流表。流表至少包括匹配域和执行域，该匹配域用于与数据包进行匹配，该执行域用于指示匹配上流表的数据包应该执行的动作。执行域包括了数据包的动作标识，例如转发、丢弃、上送 SDN 控制器等，执行域还包括了数据包的路由信息，例如数据包的目的端口标识等。

贯穿本说明书，数据流（英文全称：data flow）指示能够匹配相同流表的一系列数据包。具体的，同一数据流中的数据包的匹配信息，均可以匹配上该数据流对应的流表的匹配域。

贯穿本说明书，虚拟交换机为安装在计算设备上的，通过软件实现的交换设备，常用于 SDN 中。常见的虚拟交换机包括 Open vSwitch，缩写为 OVS，OVS 为一个开源项目提供的虚拟交换机。

贯穿本说明书，overlay 类型数据包指代采用了 overlay 封装技术处理的数据包，具体的 overlay 封装技术包括了虚拟可扩展局域网(英文全称：virtual extensible local area network，缩写：VXLAN) 技术，使用通用路由的网络虚拟化（英文全称：network virtualization using generic routing encapsulation，缩写：NVGRE）技术和无状态传输隧道（英文全称：stateless transport tunneling，缩写：STT）技术。Overlay 类型数据包包括两个部分，overlay 头和原始数据包，该原始数据包指代 VM 发出的数据包或经过虚拟交换机的端口发往 VM 的数据包，该 overlay 头叠加在该原始数据包上，以用于该 overlay 类型数据包在 overlay 网络中传输。不同的 overlay 封装技术对应于不同的 overlay 头。

贯穿本说明书，流表集合中包括了一个或多个流表。本申请中包括两个流表集合，即虚拟交换机可访问流表集合和 NIC 可访问流表集合。虚拟交换机可访问流表集合中的流表供虚拟交换机使用，NIC 可访问流表集合中的流表供 NIC 使用。具体的，如图 2c 所示，虚拟交换机可访问流表集合一般存储于计算设备

的存储设备中，NIC 可访问流表集合可以存储于计算设备的存储设备中，也可以存储于 NIC 内部的存储设备中。若虚拟交换机可访问流表集合和 NIC 可访问流表集合均存储于计算设备的存储设备中，计算设备在其存储设备中为虚拟交换机可访问流表集合和 NIC 可访问流表集合分别开辟一块内存空间。本申请的附图中，以 NIC 可访问流表集合存储于计算设备的存储设备中为例进行介绍，本领域技术人员可以直接推导出 NIC 可访问流表集合存储于 NIC 内部的情况。

贯穿本说明书，示例性的采用了 SR-IOV 的 NIC 与 VM 直连的技术，在实际使用中也可以采用其他支持 NIC 与 VM 直连的技术。

本申请实施例所应用的 SDN 架构

图 2a 为本申请实施例所应用的 SDN 架构的示意图，图 2a 中示意性的采用了集中式的 SDN 控制器，实际中 SDN 控制器也可以分布式的部署于各个计算设备。

各个计算设备上的硬件层设置有 NIC，处理器以及存储设备。本申请中，将每个计算设备除 NIC 之外的部分称之为主机。其中，处理器可以为中央处理器（英文：central processing unit，缩写：CPU），存储设备包括易失性存储器（英文：volatile memory），例如随机存取存储器（英文：random-access memory，缩写：RAM），以及非易失性存储器（英文：non-volatile memory），例如只读存储器（英文：read-only memory，缩写：ROM）、快闪存储器、HDD 或 SSD 等。每个主机运行时，其硬件层支持软件层内的虚拟交换机以及多个 VM 的运行。每个计算设备内的主机和 NIC 建立通信连接，主机通过 NIC 与外部网络通信，例如首先由 NIC 从外部网络获取发往该主机上运行的 VM 数据包，然后发送至主机上运行的 VM，而该主机上运行的 VM 发往外部网络的数据包首先发送至 NIC，然后通过 NIC 发送至外部网络。

下面以计算设备 1 为例展示现有技术中的数据包处理流程与本申请提供的数据包处理流程的区别。如图 2b，现有技术中，计算设备 1 内的 NIC 从外部网络接收到数据包后，如果判断该数据包的目的地属于计算设备 1 则将该数据包发送至虚拟交换机，则由虚拟交换机将该数据包与虚拟交换机可访问流表集合中的流表进行匹配，并根据匹配上的流表的指示，将该数据包发送至与该虚拟交换机相连的目的 VM。由以上数据包的处理流程可见，现有技术中，数据包处

理过程中主要的运行压力集中在虚拟交换机上，而虚拟交换机的运行依赖于计算设备上的硬件层的资源，虚拟交换机占用的处理器和存储设备资源越多，计算设备上能够用于 VM 运行的资源就越少，而如果限定虚拟交换机能够占用的硬件层的资源的上限，那么随着数据包流量的增大，虚拟交换机的性能将难以保证。

如图 2c，本申请提供的数据包处理流程中，计算设备 1 内的 NIC 从外部网络接收到数据包后，如果判断该数据包的目的 VM 运行于计算设备 1 上，则将该数据包与 NIC 可访问流表集合中的流表进行匹配，并根据匹配上的流表的指示，将该数据包发送至与该 NIC 相连的目的 VM。该 NIC 可访问流表集合内的流表来源于主机，NIC 如果无法将该数据包匹配上 NIC 可访问流表集合中的流表，就会将该数据包发送至虚拟交换机，虚拟交换机向 SDN 控制器请求获取该数据包对应的流表，并将获取的该数据包对应的流表发送至 NIC 可访问流表集合，以供 NIC 在接下来的数据包处理过程中使用。

由以上数据包的处理流程可见，在本申请提供的数据包处理流程中，数据包的处理过程中的一部分运行压力被转移到了 NIC 上，而 NIC 作为一个硬件设备，不仅处理效率高，并且其运行无需占用硬件层的其他资源。

需要说明的是，示意性的，图 2c 中的计算设备 1 上的所有 VM 都与 NIC 相连，实际上也可以只有部分 VM 与 NIC 相连，其他部分 VM 与虚拟交换机相连，具体 VM 的配置方式并不限定于必须全部都与 NIC 相连。

图 2a 和图 2c 中的计算设备可以通过图 3 所示的计算设备 200 实现，其组织结构示意图如图 3 所示，计算设备 200 包括了主机以及 NIC206，NIC206 通过主机的总线 208 与主机的处理器 202 以及存储器 204 建立通信连接，NIC206、处理器 202 和存储器 204 之间也可以通过无线传输等其他手段实现通信。计算设备 200 通过 NIC206 与外部网络通信。

工作状态下，主机运行了至少一个 VM 以及虚拟交换机，且用于实现图 4 提供的数据包处理方法中主机侧的方法的程序代码保存在存储设备 204 中，并由处理器 202 执行。工作状态下，NIC206 执行图 4 提供的数据包处理方法中 NIC 侧的方法。

本申请还提供了一种数据包处理方法，前述SDN架构中的计算设备运行时执行该方法，其流程示意图如图4所示。

步骤402，计算设备的主机接收第一虚拟交换机端口的配置信息，该第一虚拟交换机端口的配置信息指示在虚拟交换机上建立至少一个虚拟交换机端口，每个虚拟交换机端口对应该主机上运行的一个VM。

步骤404，该主机生成第二虚拟交换机端口的配置信息，并将该第二虚拟交换机端口的配置信息发送至计算设备的NIC。

具体的，该主机上运行的拦截模块，获取该第一虚拟交换机端口的配置信息，将该第一虚拟交换机端口的配置信息发送至该主机上运行的NIC驱动，该NIC驱动根据该第一虚拟交换机端口的配置信息，生成该第二虚拟交换机端口的配置信息，并发送至该NIC。该第一虚拟交换机端口的配置信息与该第二虚拟交换机端口的配置信息的功能类似，该NIC驱动对其转换主要为了符合NIC驱动与NIC通信的规范。

步骤406，该NIC根据该第二虚拟交换机端口的配置信息，在该NIC上配置至少一个NIC端口，每个NIC端口通过SR-I/OV技术与该主机上运行的一个VM连接。

NIC端口具体可以为SR-I/OV技术定义的虚拟功能（英文全称：virtual function，缩写：VF）的端口。

步骤402至步骤406为可选步骤，且步骤402至步骤406为该虚拟交换机和该NIC的配置过程，无须每次执行步骤408及步骤408的后续步骤前都执行一次步骤402至步骤406。如图5a或图5b或图5c，通过该配置过程，主机上运行的VM通过NIC端口与NIC连接，虚拟交换机上虽然建立了与VM一一对应的VS端口，但主机上运行的VM并不与虚拟交换机连接。

由于VS端口与VM一一对应，同时VM与NIC端口一一对应，因此VS端口与NIC端口一一对应。在步骤402至步骤406的执行过程中将VS端口与NIC端口的对应关系存入该虚拟交换机，或将VS端口与NIC端口的对应关系存入该NIC。

如图5b或图5c，该虚拟交换机和该NIC的配置过程中，或该虚拟交换机和该NIC的配置过程之前或之后，还需要配置该虚拟交换机与该NIC通信的至少一个

队列，用于该虚拟交换机将从NIC接收到的数据包返回给该NIC。队列的配置有两种方式，其一如图5b所示，该虚拟交换机与该NIC通过一个队列通信，该虚拟交换机将需要发往该NIC的全部数据包发送至该队列；其二如图5c所示，该虚拟交换机与该NIC通过n个队列通信，n为该主机上运行的VM的数量，每一个队列与一个VM对应。

该配置过程无须上层管理设备的感知，由该计算设备将本应连接于虚拟交换机的VM连接于NIC上，无须上层管理设备对配置信息进行修改，提升了配置过程的兼容性和降低了实现难度。

步骤408，该NIC接收第一overlay类型数据包，该第一overlay类型数据包包括第一overlay头和该第一数据包，该第一overlay头包括VXLAN头，或NVGRE头，或STT头。

该第一overlay类型数据包可以为外部网络发送至该NIC的。

步骤410，该NIC剥去该第一overlay类型数据包的第一overlay头，获取该第一数据包。

相对于现有技术中，由主机执行剥离overlay头的动作，由NIC剥离overlay头降低了主机的工作负担。

需要说明的是，如果步骤408中该NIC接收的不是overlay类型数据包，而是直接接收到第一数据包，则无须执行步骤410。

步骤412，该NIC根据该第一数据包的匹配信息查询NIC可访问流表集合。若无法匹配NIC可访问流表集合中的任一流表，则执行步骤414、步骤416、步骤4181，或步骤4182至步骤4184，或步骤4185至步骤4186。若能够匹配NIC可访问流表集合中的流表，则执行步骤420。

如果该第一数据包的匹配信息无法匹配NIC可访问流表集合中的任一流表，则该第一数据包为该第一数据包所在的数据流的首个数据包，或该第一数据包不是该数据流的首个数据包，但该数据流对应的流表在NIC可访问流表集合中已经被删除。

如果该第一数据包的匹配信息能够匹配NIC可访问流表集合中的流表，则说明NIC可访问流表集合中已经存有该第一数据包所在的数据流对应的流表。

步骤414，该NIC通过主机端口向该虚拟交换机转发该第一数据包。

该主机端口可以为SR-I/OV技术定义的物理功能（英文全称：physical function，缩写：PF）的端口。

步骤416，该虚拟交换机在接收到该第一数据包后，获取该数据流对应的流表，该数据流对应的流表被加入到该NIC可访问流表集合。

该虚拟交换机获取该第一数据包后，将该第一数据包发送至SDN控制器，并接收SDN控制器根据该第一数据包生成的该数据流对应的流表。虚拟交换机可访问流表集合中，还可能存储有生成该数据流对应的流表所需的信息，例如慢表（英文全称：slow table），则该虚拟交换机根据该信息生成该数据流对应的流表即可，无须将该第一数据包发送至SDN控制器。

该虚拟交换机将该数据流对应的流表存入虚拟交换机可访问流表集合和NIC可访问流表集合。或者，该主机上运行的监控模块监控该虚拟交换机获取该数据流对应的流表，该监控模块将该数据流对应的流表存入NIC可访问流表集合。

由于该数据流对应的流表由SDN控制器生成，而SDN控制器无须知道主机上运行的VM实际连接于NIC。因此该数据流的数据包的路由信息具体可以包括VS端口标识，而VS端口与VM一一对应，同时VM与NIC端口一一对应，因此VS端口与NIC端口一一对应，该数据流的数据包的路由信息指示目的VM的NIC端口。

步骤416后，将该第一数据包发送至其目的VM有三种可选的方案，分别为步骤4181、步骤4182至步骤4184和步骤4185至步骤4186，实际步骤416后可以执行这三种可选方案中的任一种。

步骤4181，该NIC根据该第一数据包的匹配信息查询该NIC可访问流表集合，获取该数据流对应的流表，并根据该数据流的数据包的路由信息转发该第一数据包至该目的VM。

这种情况下，该NIC需要存有VS端口标识与NIC端口标识的对应关系，该NIC获取该数据流的数据包的路由信息包括的VS端口标识后，将VS端口标识转换为NIC端口标识，并将该第一数据包从该NIC端口标识对应的NIC端口发出。

步骤416中该虚拟交换机或该监控模块将该数据流对应的流表存储到NIC可

访问流表集合后，向该NIC发送通知消息，该通知消息用于通知该NIC该数据流对应的流表已经存储于NIC可访问流表集合。该NIC接收到该通知信息后，根据该第一数据包的匹配信息在该NIC可访问流表集合中可以匹配上该数据流对应的流表。

或者，该NIC在步骤414后，周期性的根据该第一数据包的匹配信息匹配该NIC可访问流表集合的流表，在步骤416执行完毕后该NIC的下一次匹配中，该NIC根据该第一数据包的匹配信息在该NIC可访问流表集合中可以匹配上该数据流对应的流表。

可选步骤4181无须虚拟交换机将该第一数据包和该数据流对应的流表进行匹配，降低了虚拟交换机的工作负担。

步骤4182，该虚拟交换机将该第一数据包与该虚拟交换机可访问流表集合中的该数据流对应的流表进行匹配，获取该数据流的数据包的路由信息。

步骤4183，该虚拟交换机，根据该数据流的数据包的路由信息生成返回的第一数据包，将该返回的第一数据包发送至该NIC，该返回的第一数据包中包含对应于该目的VM的端口标识，该端口标识为虚拟交换机端口标识或NIC端口标识。

步骤4184，该NIC接收该返回的第一数据包，并根据该端口标识，转发该第二数据包到该目的VM。

如图5b，如步骤416中所述，该数据流的数据包的路由信息具体可以包括VS端口标识。例如，该第一数据包的目的VM为VM-1，VM-1在虚拟交换机上对应的端口为VS端口1，VM-1在NIC上对应的端口为NIC端口1，则该数据流的数据包的路由信息包括VS端口1。步骤4183中，该虚拟交换机生成的该返回的第一数据包中包含了该第一数据包的目的VM的端口标识和该第一数据包，而该第一数据包的目的VM的端口标识为VS端口1或NIC端口1。

步骤4183中可选的，该虚拟交换机将该数据流的数据包的路由信息添加到该返回的第一数据包中，则该目的VM的端口标识为VS端口1，并通过队列发送至该NIC。这种情况下，该NIC需要存有VS端口标识与NIC端口标识的对应关系，该NIC接收到该返回的第一数据包后，将该VS端口1转换为NIC端口1，并将该第

一数据包通过NIC端口1发送至VM-1。这种实施方式下该虚拟交换机的负载更低，提升了主机的工作效率。

步骤4183中可选的，该虚拟交换机获取将该数据流的数据包的路由信息后，将该数据流的数据包的路由信息包括的VS端口1转换为NIC端口1，并将NIC端口1添加到该返回的第一数据包中，则该目的VM的端口标识为NIC端口1，并通过队列将该返回的第一数据包发送至该NIC。这种情况下，该虚拟交换机需要存有VS端口标识与NIC端口标识的对应关系。该NIC接收到该返回的第一数据包后，将该第一数据包通过NIC端口1发送至VM-1。这种实施方式下，NIC无须对端口标识进行转换，可以更高效的处理数据包。

步骤4185，该虚拟交换机将该第一数据包发送至该目的VM对应的队列。该虚拟交换机与该NIC通过至少一个队列通信，每个主机上运行的VM对应一个队列。

步骤4186，该NIC从该目的VM对应的队列接收该第一数据包，并该NIC根据该目的VM对应的队列的队列信息，从该目的VM对应的NIC端口转发该第一数据包至该目的VM。

如图5c，该虚拟交换机与该NIC通过至少n个队列通信，n为该主机上运行的VM的数量，每一个队列与一个VM对应。该虚拟交换机将该第一数据包与该虚拟交换机可访问流表集合中的该数据流对应的流表进行匹配，获取该数据流的数据包的路由信息后，例如为VS端口1，而VS端口1对应VM1且VM1对应于队列1，则该虚拟机交换机将该第一数据包发送至队列1。

该NIC从队列1获取该第一数据包，该主机上运行的NIC驱动向该NIC发送队列信息，该队列信息用于通知该NIC该第一数据包来自队列1。由于队列1与VM1的对应，同时VM1与NIC端口1对应，则该NIC通过NIC端口1将该第一数据包发送至VM1。此种方式要求NIC存储有队列与NIC端口的对应关系。这种实施方式与前述两种可选方案相比，虚拟交换机和NIC均无需将该数据流的数据包的路由信息转换为NIC端口标识，提升了数据包的转发效率。

实际的SDN中，VM一般设有安全组，因此在这三种可选方案中确认了该第一数据包的目的VM后，可选的，还需要确认该第一数据包通过安全组检查后，

才将该第一数据包发送至该第一数据包的目的VM。

可选的，若设置有静态安全组，首先确定该第一数据包的目的VM是否属于某一静态安全组，若确定该第一数据包的目的VM属于某一静态安全组，则判断该第一数据包能否匹配上该静态安全组的任一规则，若该第一数据包能够匹配上该静态安全组的至少一条规则，则该第一数据包通过静态安全组检查。若该第一数据包的目的VM不属于任一静态安全组，则无需对该第一数据包进行静态安全组检查，根据第一预设规则直接对该第一数据包进行处理，例如将该第一数据包发送至该第一数据包的目的VM。若该第一数据包的目的VM属于某一静态安全组，但该第一数据包无法匹配该静态安全组的任一规则，则该第一数据包无法通过安全组检查，根据第二预设规则处理该第一数据包，例如丢弃该第一数据包。

以上为设置有白名单的静态安全组的场景，如果设置的为黑名单的静态安全组，与白名单的场景相反，如果该第一数据包属于某一静态安全组但无法匹配该静态安全组的任一规则，则该第一数据包通过静态安全组检查。而如果该第一数据包的目的VM属于某一静态安全组，且该第一数据包可以匹配该静态安全组的至少一条规则，则该第一数据包无法通过静态安全组检查。

可选的，若设置有动态安全组，则首先判断该第一数据包的目的VM是否属于动态安全组，如果该目的VM属于动态安全组，则根据该第一数据包查询连接跟踪表（英文全称：connection track table），确认该第一数据包属于哪个连接，并确定该第一数据包该连接的状态以及该第一数据包对应的处理动作，如果该第一数据包的处理动作指示转发该第一数据包至该第一数据包的目的VM，则该第一数据包通过动态安全组检查。

以上静态安全组和动态安全组可以同时设置，此时通过静态安全组检查和动态安全组检查的数据包才通过了安全组检查。步骤416后采用步骤4182至步骤4184或步骤4185至步骤4186的情况下，安全组检查可以由主机上运行的安全组模块实现，因此如果安全组模块确认该第一数据包无法通过安全组检查，则无需将该第一数据包发送至该NIC，提升了NIC的工作效率。而步骤416后采用步骤4181的情况下，对于发送至虚拟交换机的数据包的安全组检查可以在数据包发送

回NIC后，由NIC执行。

步骤420，该NIC根据匹配上的流表中包括的路由信息，转发该第一数据包到该第一数据包的目的VM。

该NIC根据该第一数据包的匹配信息，在该NIC可访问流表集合中匹配上该第一数据包所在数据流对应的流表，根据该流表包括的该数据流的数据包的路由信息转发该第一数据包到该第一数据包的目的VM。

由于步骤412中，该NIC能够将该第一数据包匹配上NIC可访问流表集合中的流表，因此该第一数据包不是所在数据流的首个数据包。

步骤420中，该NIC根据匹配上的流表中包括的路由信息后，参照前述安全组检查过程，可选的，该NIC确认该第一数据包通过安全组检查后，才将该第一数据包发送至该第一数据包的目的VM。

步骤4181或步骤4184或步骤4186或步骤420之后，该NIC继续接收该数据流的后续数据包的情况下，例如第二overlay数据包，该第二overlay数据包包括第二数据包和第二数据包对应的第二overlay头，如果该数据流对应的流表仍存储于NIC可访问流表集合中，则NIC根据该数据流的数据包的路由信息转发该第二数据包到该目的VM。而实际中由于NIC可访问流表集合中的流表可能随着时间更新，因此虽然在步骤416中该数据流对应的流表被加入到该NIC可访问流表集合，但NIC根据该第二数据包的匹配信息无法匹配NIC可访问流表集合中的任一流表，这种情况下对该后续数据包执行步骤414、步骤416、步骤4181或步骤4182至步骤4184或步骤4185至步骤4186。

该数据包处理方法，将数据包与流表的匹配功能卸载至NIC上执行，降低了虚拟交换机的工作负担，使得主机的硬件层资源可以更好的服务于VM，提升了计算设备的工作效率。

参考图5d，为本申请提供的另一计算设备的结构示意图，与图5a、图5b或图5c不同，该计算设备中VM-1至VM-n连接于NIC，VM-n+1至VM-n+m连接于虚拟交换机。VM-n+1至VM-n+m可以在执行图4中虚拟交换机和该NIC的配置过程之前就已经配置完毕，或在图4的虚拟交换机和该NIC的配置过程中，有选择的将

VM-n+1至VM-n+m连接于虚拟交换机，将VM-1至VM-n连接于NIC，具体可以依据主机的负载情况或根据主机接收到的配置信息中携带的信息来配置一部分VM连接于NIC，其余部分VM连接于虚拟交换机。

图5d的场景下，NIC接收到外部网络发来的数据包的情况下，如果该数据包的目的VM连接于该NIC，则对该数据包执行前述图4中的数据包处理方法，如果该数据包的目的VM连接于该虚拟交换机，则该NIC直接将该数据包发送至该虚拟交换机。由该虚拟交换机完成该数据包的流表匹配，并将该数据包发送至其目的VM。具体的，可以通过流表的设置使得该NIC实现上述功能。例如，目的VM连接于该虚拟交换机的数据包所在数据流对应的流表，不会被存入NIC可访问流表集合，则该NIC接收到目的VM连接于该虚拟交换机的数据包时，无法在NIC可访问流表集合中匹配上流表，则会将该数据包发送至该虚拟交换机；或修改目的VM连接于该虚拟交换机的数据包所在数据流对应的流表后，再将该流表存入NIC可访问流表集合，对该流表的修改包括将该流表的路由信息修改为主机端口，则该NIC接收到目的VM连接于该虚拟交换机的数据包时，在NIC可访问流表集合中匹配上的流表指示将该数据包通过主机端口发送至该虚拟交换机。

本申请还提供了一种NIC600，该NIC600可以为前述任一附图提供的NIC。该NIC600的组织结构示意图如图6a所示，包括主机接口602，网络接口604以及处理芯片606。网络接口604用于与外部网络通信且网络接口604与处理芯片606建立通信连接。主机接口602用于与NIC600所连接的主机上运行的虚拟交换机、VM、NIC驱动等通信且主机接口602与处理芯片606建立通信连接。前述数据包处理方法中，NIC上建立的NIC端口与主机端口为虚拟端口，主机端口与NIC端口实际通过主机接口602实现与主机的通信。参考图3，主机接口602实际可以为NIC600与计算设备的总线连接的接口。

主机接口602，用于从该NIC连接的主机获取虚拟交换机端口的配置信息。将该虚拟交换机端口的配置信息发送至处理芯片606。

处理芯片606，用于根据该虚拟交换机端口的配置信息将NIC600与该主机上运行的VM连接。

具体的，参考前述步骤406，处理芯片606根据该虚拟交换机端口的配置信息在该NIC上配置至少一个NIC端口，每个NIC端口对应该主机上运行的一个VM。

以上为NIC600的配置过程中NIC600中各个单元的功能，NIC600还可以用于数据包的处理，参考前述数据包处理方法中的步骤408及其步骤408之后的步骤。

网络接口604，还用于接收第一overlay类型数据包，参考步骤408。

网络接口604将接收到的该第一overlay类型数据包发送至处理芯片606。

处理芯片606，还用于接收到该第一overlay类型数据包后，对该第一overlay类型数据包的处理参考前述步骤410、步骤412。

处理芯片606，还用于执行步骤412。在步骤412中确定该第一数据包无法匹配NIC可访问流表集合中的任一流表的情况下，执行步骤414及后续步骤。在步骤412中确定该第一数据包能够匹配NIC可访问流表集合中的流表的情况下，执行步骤420。

步骤414后，处理芯片606执行三种可选方案之任一，这三种可选方案分别对应于前述步骤4181、步骤4182至步骤4184中NIC侧执行的部分、步骤4185至步骤4186中NIC侧执行的部分。

可选方案一，对应前述步骤4181，处理芯片606根据该第一数据包的匹配信息查询该NIC可访问流表集合，获取该第一数据包所在的数据流对应的流表，也即获取该数据流的数据包的路由信息包括的VS端口标识，将VS端口标识转换为NIC端口标识。

可选方案二，对应前述步骤4182至步骤4184中NIC侧执行的部分。处理芯片606接收该返回的第一数据包，处理芯片606根据该返回的第一数据包中携带的端口标识获取NIC端口标识。如果该端口标识为虚拟交换机端口标识，则处理芯片606将该虚拟交换机端口标识转换为NIC端口标识。该端口标识也可能为NIC端口标识。

可选方案三，对应前述步骤4185至步骤4186中NIC侧执行的部分。处理芯片606从n个队列中的一个队列接收该第一数据包，由于处理芯片606预先配置有每个队列与NIC端口的对应关系，因此处理芯片606可以获得接收该第一数据包的队列对应的NIC端口标识。

处理芯片606执行步骤412，并确定该第一数据包能够匹配NIC可访问流表集合中的流表的情况下，获取匹配上的流表中包括的该数据流的数据包的路由信息。该数据流的数据包的路由信息可以包括VS端口标识，处理芯片606将VS端口标识转换为NIC端口标识。

处理芯片606执行步骤412后，无论在该第一数据包能否匹配NIC可访问流表集合中的流表的情况下，均会获取该第一数据包对应的NIC端口标识，该NIC端口标识对应于该第一数据包所在数据流的目的VM。处理芯片606确定该第一数据包的目的VM后，还用于对该第一数据包进行安全组检查。处理芯片606在确定该第一数据包的安全组检查通过之后，通过主机接口602将该第一数据包到该目的VM。具体的处理芯片606对该第一数据包进行安全组检查的过程，参考前述数据包处理方法。

需要说明的是，如果网络接口604接收的不是overlay类型数据包，则处理芯片606无须执行步骤410。

以上提供的NIC实现了流表匹配功能，对于在NIC可访问流表集合中的数据包，将无需发送至虚拟交换机进行处理，降低了主机的负荷，提升了与该NIC相连的主机的工作效率。

处理芯片606可以通过专用集成电路（英文：application-specific integrated circuit，缩写：ASIC）实现，或可编程逻辑器件（英文：programmable logic device，缩写：PLD）实现。上述PLD可以是复杂可编程逻辑器件（英文：complex programmable logic device，缩写：CPLD），现场可编程门阵列（英文：field programmable gate array，缩写：FPGA），通用阵列逻辑（英文：generic array logic，缩写：GAL）或其任意组合。

具体的，如图6b所示，处理芯片606可以包括overlay芯片6062、流表匹配芯片6064以及安全组检查芯片6066。其中，overlay芯片6062用于将网络接口804发送至处理芯片606的overlay类型数据包的overlay头剥离。流表匹配芯片6064用于将剥离了overlay头的数据包与存储于NIC可访问流表集合中的流表进行匹配。安全组检查芯片6066用于确定在流表匹配芯片8064中匹配上流表的数据包是否通过安

全组检查，并将通过安全组检查的数据包通过主机接口602发送至目的VM。Overlay芯片6062与安全组检查芯片6066为可选组件。

处理芯片606还可以通过处理器、存储设备以及逻辑芯片实现，该逻辑芯片可以由PLD或ASIC实现。该处理芯片606运行时，该处理器和该逻辑芯片各执行一部分功能，两者功能的分配可以有多种。示例性的，如图6c所示，逻辑芯片用于将网络接口604发送至处理芯片606的overlay类型数据包的overlay头剥离。处理芯片606内的处理器工作时，读取存储器内的代码，用于读取NIC可访问流表集合中的流表，并将NIC可访问流表集合中的流表发送至逻辑芯片，以供逻辑芯片将剥离了overlay头的数据包与该流表进行匹配。处理器还用于读取进行安全组检查所需的信息，并将安全组检查所需的信息发送至逻辑芯片，以供逻辑芯片对该数据包进行安全组检查。

在图6c所示的NIC600的处理芯片606中，逻辑芯片也可以由overlay子芯片、流表匹配子芯片、安全组检查子芯片构成。overlay子芯片、安全组检查子芯片为可选组件。overlay子芯片用于将overlay类型数据包的overlay头剥离。图6c所示的NIC600的处理芯片606中处理器用于获取流表匹配或安全组检查所需的信息并将流表匹配或安全组检查所需的信息发送至该逻辑芯片，而流表匹配子芯片根据流表匹配所需的信息完成数据包的流表匹配，安全组检查子芯片根据安全组检查所需的信息完成数据包的安全组检查。

本申请还提供了一种数据包处理方法，前述任一附图中的NIC运行时执行该方法。该方法具体参考图4对应的数据包处理方法中NIC侧执行的部分。

本申请还提供了一种配置方法，前述任一附图中的主机运行时执行该方法。该方法具体参考图4对应的数据包处理方法中的步骤402和步骤404。

本申请还提供了一种数据包处理方法，前述任一附图中主机运行时执行该方法。该方法具体参考图4对应的数据包处理方法中的步骤408后主机侧执行的方法。具体包括，步骤416、步骤4181中将该数据流对应的流表存储到NIC可访问流表集合后向该NIC发送通知消息的部分，或步骤4182和步骤4183，或步骤4185。

在上述实施例中，对各个实施例的描述都各有侧重，某个实施例中没有详述的部分，可以参见其他实施例的相关描述。

结合本申请公开内容所描述的方法可以由处理器执行软件指令的方式来实现。软件指令可以由相应的软件模块组成，软件模块可以被存放于 RAM、快闪存储器、ROM、可擦除可编程只读存储器（英文：erasable programmable read only memory，缩写：EPROM）、电可擦可编程只读存储器（英文：electrically erasable programmable read only memory，缩写：EEPROM）、硬盘、光盘或者本领域熟知的任何其它形式的存储介质中。

本领域技术人员应该可以意识到，在上述一个或多个示例中，本申请所描述的功能可以用硬件或软件来实现。当使用软件实现时，可以将这些功能存储在计算机可读介质中或者作为计算机可读介质上的一个或多个指令或代码进行传输。存储介质可以是通用或专用计算机能够存取的任何可用介质。

以上该的具体实施方式，对本申请的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上该仅为本申请的具体实施方式而已，并不用于限定本申请的保护范围，凡在本申请的技术方案的基础之上，所做的任何修改、改进等，均应包括在本申请的保护范围之内。

权 利 要 求 书

1、一种数据包处理方法，其特征在于，所述方法应用于计算设备，所述计算设备包括网络接口卡NIC和主机，所述NIC与所述主机建立通信连接，所述主机运行虚拟机VM，所述方法包括：

所述NIC接收数据流的第一数据包；

所述NIC根据所述第一数据包的匹配信息查询流表集合；

在无法匹配到所述数据流对应的流表情况下，所述NIC向所述主机上运行的虚拟交换机转发所述第一数据包；

其中，所述虚拟交换机在接收到所述第一数据包后，从软件定义网络SDN控制器获取所述数据流对应的流表，以便于所述数据流对应的流表被加入到所述流表集合。

2、如权利要求1所述的数据包处理方法，其特征在于，所述主机上运行的虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应所述主机上运行的一个VM，所述NIC接收所述第一数据包前，所述方法还包括：

所述NIC接收所述虚拟交换机端口的配置信息；

所述NIC根据所述虚拟交换机端口的配置信息，在所述NIC上配置至少一个NIC端口，每个NIC端口通过单根-输入/输出虚拟化SR-I/OV技术与所述主机上运行的一个VM连接。

3、如权利要求2所述的数据包处理方法，其特征在于，所述数据流对应的流表包括所述数据流的数据包的路由信息，在向所述虚拟交换机转发所述第一数据包之后，所述方法还包括：

所述NIC根据所述第一数据包的匹配信息查询所述流表集合，获取所述数据流对应的流表，并根据所述数据流的数据包的路由信息转发所述第一数据包至目的VM。

4、如权利要求2所述的数据包处理方法，其特征在于，所述数据流对应的流表包括所述数据流的数据包的路由信息，在向所述虚拟交换机转发所述第一数据包之后，所述方法还包括：

所述NIC接收所述虚拟交换机返回的第一数据包，所述返回的第一数据包中

包含对应于目的VM的端口标识，所述端口标识为虚拟交换机端口标识或NIC端口标识，所述端口标识由所述虚拟交换机根据所述数据流的数据包的路由信息添加；

所述NIC根据所述端口标识，转发所述第一数据包到所述目的VM。

5、如权利要求2所述的数据包处理方法，其特征在于，所述虚拟交换机与所述NIC通过至少一个队列通信，每个队列对应所述主机上运行的一个VM；

在向所述虚拟交换机转发所述第一数据包之后，所述方法还包括：

所述NIC从所述第一数据包的目的VM对应的队列接收所述第一数据包；

所述NIC根据所述目的VM对应的队列的队列信息，从所述目的VM对应的NIC端口转发所述第一数据包至所述目的VM。

6、一种网络接口卡NIC，其特征在于，所述NIC包括：主机接口、网络接口和处理芯片，所述网络接口用于与外部网络通信且所述网络接口与所述处理芯片建立通信连接，所述主机接口用于与主机通信且所述主机接口与所述处理芯片建立通信连接，所述主机运行VM；

所述网络接口，用于接收数据流的第一数据包；

所述处理芯片，用于根据所述第一数据包的匹配信息查询流表集合，在无法匹配到所述数据流对应的流表情况下，通过所述主机接口向所述主机上运行的虚拟交换机转发所述第一数据包；

其中，所述虚拟交换机在接收到所述第一数据包后，从软件定义网络SDN控制器获取所述数据流对应的流表，以便于所述数据流对应的流表被加入到所述流表集合。

7、如权利要求6所述的NIC，其特征在于，所述主机上运行的虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应所述主机上运行的一个VM；

所述处理芯片，还用于接收所述虚拟交换机端口的配置信息，根据所述虚拟交换机端口的配置信息，在所述NIC上配置至少一个NIC端口，每个NIC端口

通过单根-输入/输出虚拟化SR-I/OV技术与所述主机上运行的一个VM连接。

8、如权利要求7所述的NIC，其特征在于，所述处理芯片，还用于根据所述第一数据包的匹配信息查询所述流表集合，获取所述数据流对应的流表，所述数据流对应的流表包括所述数据流的数据包的路由信息，并根据所述数据流的数据包的路由信息转发所述第一数据包至目的VM。

9、如权利要求7所述的NIC，其特征在于，所述处理芯片，还用于接收所述虚拟交换机返回的第一数据包，所述返回的第一数据包中包含对应于目的VM的端口标识，所述端口标识为虚拟交换机端口标识或NIC端口标识，所述端口标识由所述虚拟交换机根据所述数据流的数据包的路由信息添加，所述数据流对应的流表包括所述数据流的数据包的路由信息；还用于根据所述端口标识，转发所述第一数据包到所述目的VM。

10、如权利要求7所述的NIC，其特征在于，所述虚拟交换机与所述NIC通过至少一个队列通信，每个队列对应所述主机上运行的一个VM；

所述处理芯片，还用于从所述第一数据包的目的VM对应的队列接收所述第二数据包，并根据所述目的VM对应的队列的队列信息，从所述目的VM对应的NIC端口转发所述第二数据包至所述目的VM。

11、一种数据包处理方法，其特征在于，所述方法应用于计算设备，所述计算设备包括网络接口卡NIC和主机，所述NIC与所述主机建立通信连接，所述主机运行虚拟机VM，所述方法包括：

所述NIC接收数据流的第二数据包；

所述NIC根据所述第二数据包的匹配信息查询流表集合，获取所述数据流对应的流表，所述数据流对应的流表包括所述数据流的数据包的路由信息；

所述NIC根据所述数据流的数据包的路由信息，转发所述第二数据包到目的VM。

12、如权利要求11所述的数据包处理方法，其特征在于，所述NIC通过单根-输入/输出虚拟化SR-I/OV技术与所述主机上运行的VM连接；

所述NIC根据所述数据流的数据包的路由信息，通过所述NIC与所述目的

VM的连接转发所述第二数据包到所述目的VM。

13、如权利要求12所述的数据包处理方法，所述主机上运行的虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应所述主机上运行的一个VM；

所述NIC接收数据流的第二数据包前，所述方法还包括：

所述NIC接收所述虚拟交换机端口的配置信息；

根据所述虚拟交换机端口的配置信息，在所述NIC上配置至少一个NIC端口，每个NIC端口通过单根-输入/输出虚拟化SR-I/OV技术与所述主机上运行的一个VM连接，所述数据流的数据包的路由信息指示所述目的VM对应的NIC端口。

14、如权利要求13所述的数据包处理方法，其特征在于，所述NIC接收数据流的第二数据包前，所述方法还包括：

所述NIC接收所述数据流的第三数据包；

所述NIC根据所述第三数据包的匹配信息查询所述流表集合；

在无法匹配到所述数据流对应的流表情况下，所述NIC向所述主机上运行的虚拟交换机转发所述第三数据包；

其中，所述虚拟交换机在接收到所述第三数据包后，从软件定义网络SDN控制器获取所述数据流对应的流表，以便于所述数据流对应的流表被加入到所述流表集合。

15、如权利要求14所述的数据包处理方法，其特征在于，在向所述虚拟交换机转发所述第三数据包之后，所述方法还包括：

所述NIC根据所述第三数据包的匹配信息查询所述流表集合，获取所述数据流对应的流表，并根据所述数据流的数据包的路由信息转发所述第三数据包至所述目的VM。

16、如权利要求14所述的数据包处理方法，其特征在于，在向所述虚拟交换机转发所述第三数据包之后，所述方法还包括：

所述NIC接收所述虚拟交换机返回的第三数据包，所述返回的第三数据包中包含对应于所述目的VM的端口标识，所述端口标识为虚拟交换机端口标识或

NIC端口标识，所述端口标识由所述虚拟交换机根据所述数据流的数据包的路由信息添加；

所述NIC根据所述端口标识，转发所述第三数据包到所述目的VM。

17、如权利要求14所述的数据包处理方法，其特征在于，所述虚拟交换机与所述NIC通过至少一个队列通信，每个队列对应所述主机上运行的一个VM；

在向所述虚拟交换机转发所述第三数据包之后，所述方法还包括：

所述NIC从所述目的VM对应的队列接收所述第三数据包；

所述NIC根据所述目的VM对应的队列的队列信息，从所述目的VM对应的NIC端口转发所述第三数据包至所述目的VM。

18、如权利要求11至17任一所述的数据包处理方法，其特征在于，所述NIC接收数据流的第二数据包包括：

所述NIC接收overlay类型数据包，所述overlay类型数据包包括overlay头和所述第二数据包，所述overlay头包括虚拟可扩展局域网VXLAN头，或使用通用路由的网络虚拟化NVGRE头，或无状态传输隧道STT头；

所述NIC剥去所述overlay类型数据包的overlay头，获取所述第二数据包。

19、如权利要求11至18任一所述的数据包处理方法，其特征在于，所述NIC转发所述第二数据包到目的VM之前，所述方法还包括：

所述NIC对所述第二数据包进行安全组检查，在确定所述第二数据包的安全组检查通过之后，执行转发所述第二数据包到所述目的VM的步骤。

20、一种网络接口卡NIC，其特征在于，所述NIC包括：主机接口、网络接口和处理芯片，所述网络接口用于与外部网络通信且所述网络接口与所述处理芯片建立通信连接，所述主机接口用于与主机通信且所述主机接口与所述处理芯片建立通信连接，所述主机运行VM；

所述网络接口，用于接收数据流的第二数据包；

所述处理芯片，用于根据所述第二数据包的匹配信息查询流表集合，获取所述数据流对应的流表，所述数据流对应的流表包括所述数据流的数据包的路

由信息；以及根据所述数据流的数据包的路由信息，转发所述第二数据包到目的VM。

21、如权利要求20所述的NIC，其特征在于，所述NIC通过单根-输入/输出虚拟化SR-I/OV技术与所述主机上运行的VM连接；

所述处理芯片，用于根据所述数据流的数据包的路由信息，通过所述NIC与所述目的VM的连接转发所述第二数据包到所述目的VM。

22、如权利要求21所述的NIC，其特征在于，所述主机上运行的虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应所述主机上运行的一个VM；

所述处理芯片，还用于接收所述虚拟交换机端口的配置信息，根据所述虚拟交换机端口的配置信息在所述NIC上配置至少一个NIC端口，每个NIC端口通过单根-输入/输出虚拟化SR-I/OV技术与所述主机上运行的一个VM连接，所述数据流的数据包的路由信息指示所述目的VM对应的NIC端口。

23、如权利要求22所述的NIC，其特征在于，所述网络接口，还用于接收所述数据流的第三数据包；

所述处理芯片，还用于根据所述第三数据包的匹配信息查询所述流表集合，在无法匹配到所述数据流对应的流表情况下，向所述主机上运行的虚拟交换机转发所述第三数据包；

其中，所述虚拟交换机在接收到所述第三数据包后，从软件定义网络SDN控制器获取所述数据流对应的流表，以便于所述数据流对应的流表被加入到所述流表集合。

24、如权利要求23所述的NIC，其特征在于，所述处理芯片，还用于根据所述第三数据包的匹配信息查询所述流表集合，获取所述数据流对应的流表，并根据所述数据流的数据包的路由信息转发所述第三数据包至所述目的VM。

25、如权利要求23所述的NIC，其特征在于，所述处理芯片，还用于接收所述虚拟交换机返回的第三数据包，所述返回的第三数据包中包含对应于所述目的VM的端口标识，所述端口标识为虚拟交换机端口标识或NIC端口标识，所述端口标识由所述虚拟交换机根据所述数据流的数据包的路由信息添加；还用于

根据所述端口标识，转发所述第三数据包到所述目的VM。

26、如权利要求23所述的NIC，其特征在于，所述虚拟交换机与所述NIC通过至少一个队列通信，每个队列对应所述主机上运行的一个VM；

所述处理芯片，还用于从所述目的VM对应的队列接收所述第三数据包，并根据所述目的VM对应的队列的队列信息，从所述目的VM对应的NIC端口转发所述第三数据包至所述目的VM。

27、如权利要求20至26任一所述的NIC，其特征在于，所述网络接口，用于接收overlay类型数据包，所述overlay类型数据包包括overlay头和所述第二数据包，所述overlay头包括虚拟可扩展局域网VXLAN头，或使用通用路由的网络虚拟化NVGRE头，或无状态传输隧道STT头；

所述处理芯片，用于剥去所述overlay类型数据包的overlay头，获取所述第二数据包。

28、如权利要求20至27任一所述的NIC，其特征在于，所述处理芯片转发所述第二数据包到所述目的VM之前，还用于对所述第二数据包进行安全组检查，在确定所述第二数据包的安全组检查通过之后，执行转发所述第二数据包到所述目的VM的步骤。

29、一种计算设备，其特征在于，所述计算设备包括网络接口卡NIC和主机，所述NIC与所述主机建立通信连接，所述主机上运行虚拟机VM和虚拟交换机，所述虚拟交换机上配置至少一个虚拟交换机端口，每个虚拟交换机端口对应所述主机上运行的一个VM；

所述主机，用于向所述NIC发送所述虚拟交换机端口的配置信息；

所述NIC，用于根据所述虚拟交换机端口的配置信息，在所述NIC上配置至少一个NIC端口，每个NIC端口通过单根-输入/输出虚拟化SR-I/OV技术与所述主机上运行的一个VM连接；

所述NIC，还用于接收数据流的第二数据包，根据所述第二数据包的匹配信息查询流表集合，获取所述数据流对应的流表，所述数据流对应的流表包括所述数据流的数据包的路由信息，所述数据流的数据包的路由信息指示目的VM对

应的NIC端口，以及根据所述数据流的数据包的路由信息，转发所述第二数据包到所述目的VM。

30、如权利要求29所述的计算设备，其特征在于，所述NIC，还用于接收所述数据流的第三数据包，根据所述第三数据包的匹配信息查询所述流表集合，在无法匹配到所述数据流对应的流表情况下，向所述主机转发所述第三数据包；

所述主机，用于在接收到所述第三数据包后，从软件定义网络SDN控制器获取所述数据流对应的流表，以便于所述数据流对应的流表被加入到所述流表集合。

31、如权利要求30所述的计算设备，其特征在于，所述NIC，还用于根据所述第三数据包的匹配信息查询所述流表集合，获取所述数据流对应的流表，以及根据所述数据流的数据包的路由信息转发所述第三数据包至所述目的VM。

32、如权利要求30所述的计算设备，其特征在于，所述主机，还用于生成返回的第三数据包，所述返回的第三数据包中包含对应于所述目的VM的端口标识，所述端口标识为虚拟交换机端口标识或NIC端口标识，所述端口标识由所述主机根据所述数据流的数据包的路由信息添加；

所述NIC，还用于接收所述返回的第三数据包，并根据所述端口标识，转发所述第三数据包到所述目的VM。

33、如权利要求30所述的计算设备，其特征在于，所述虚拟交换机与所述NIC通过至少一个队列通信，每个队列对应所述主机上运行的一个VM；

所述主机，还用于将所述第三数据包发送至所述目的VM对应的队列；

所述NIC，还用于从所述目的VM对应的队列接收所述第三数据包，以及根据所述目的VM对应的队列的队列信息，从所述目的VM对应的NIC端口转发所述第三数据包至所述目的VM。

34、如权利要求29至33任一所述的计算设备，其特征在于，所述NIC，具体用于接收overlay类型数据包，所述overlay类型数据包包括overlay头和所述第二数据包，所述overlay头包括虚拟可扩展局域网VXLAN头，或使用通用路由的网络虚拟化NVGRE头，或无状态传输隧道STT头；还用于剥去所述overlay类型数据包的overlay头，获取所述第二数据包。

35、如权利要求29至34任一所述的计算设备，其特征在于，所述NIC转发所述第二数据包到目的VM之前，还用于对所述第二数据包进行安全组检查，在确定所述第二数据包的安全组检查通过之后，执行转发所述第二数据包到所述目的VM的步骤。

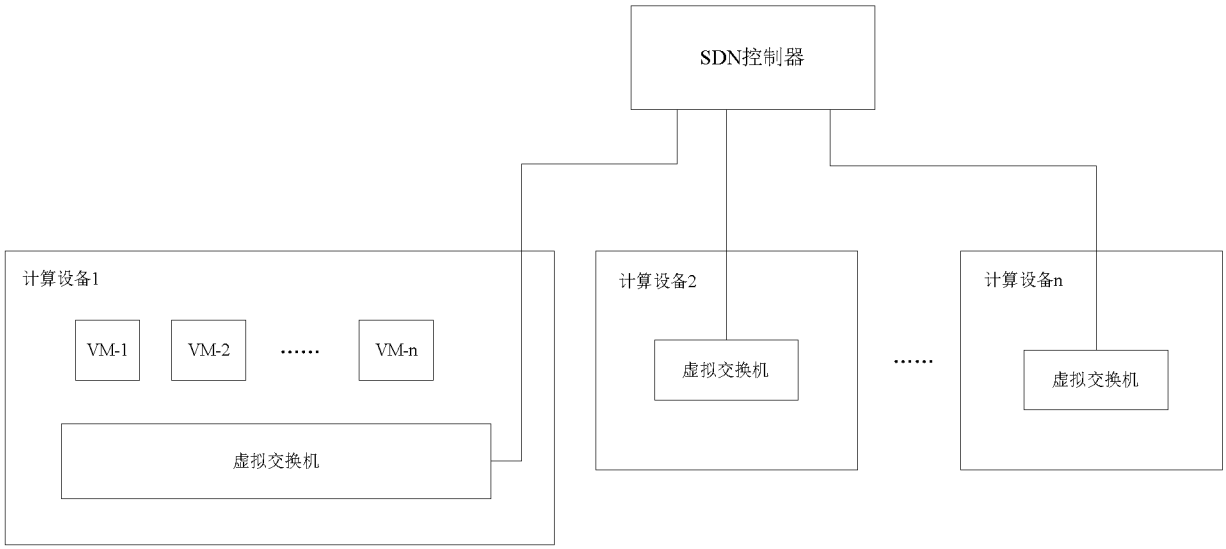
SDN控制器
计算设备1
VM-1
VM-2
……
VM-n
虚拟交换机
计算设备2
虚拟交换机
……
计算设备n
虚拟交换机

图 1

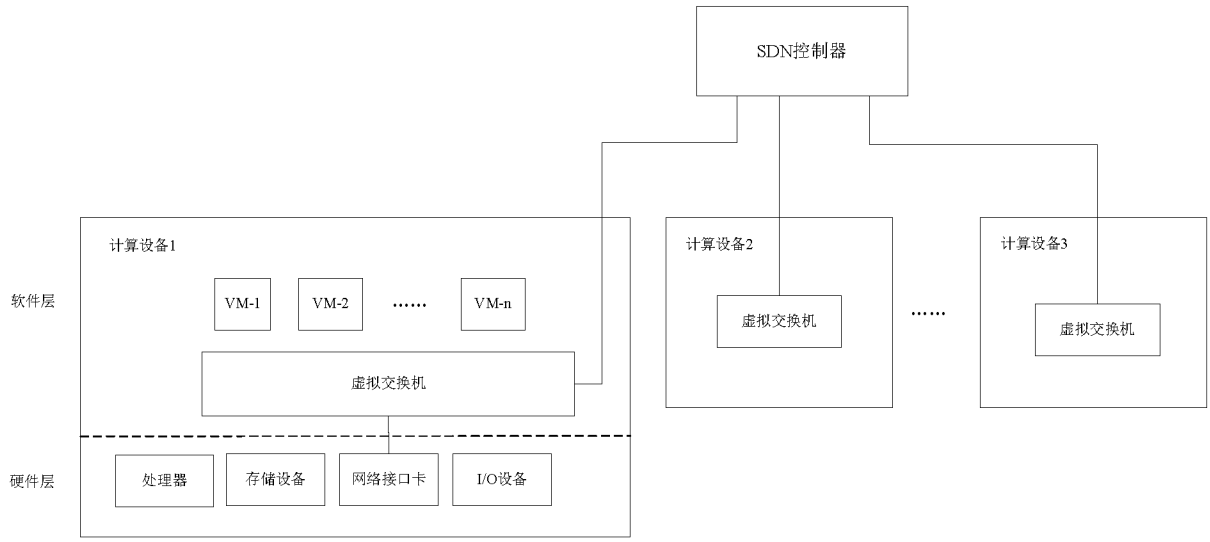
SDN控制器
计算设备1
软件层
VM-1
VM-2
……
VM-n
虚拟交换机
硬件层
处理器
存储设备
网络接口卡
I/O设备
计算设备2
虚拟交换机
……
计算设备3
虚拟交换机

图 2a

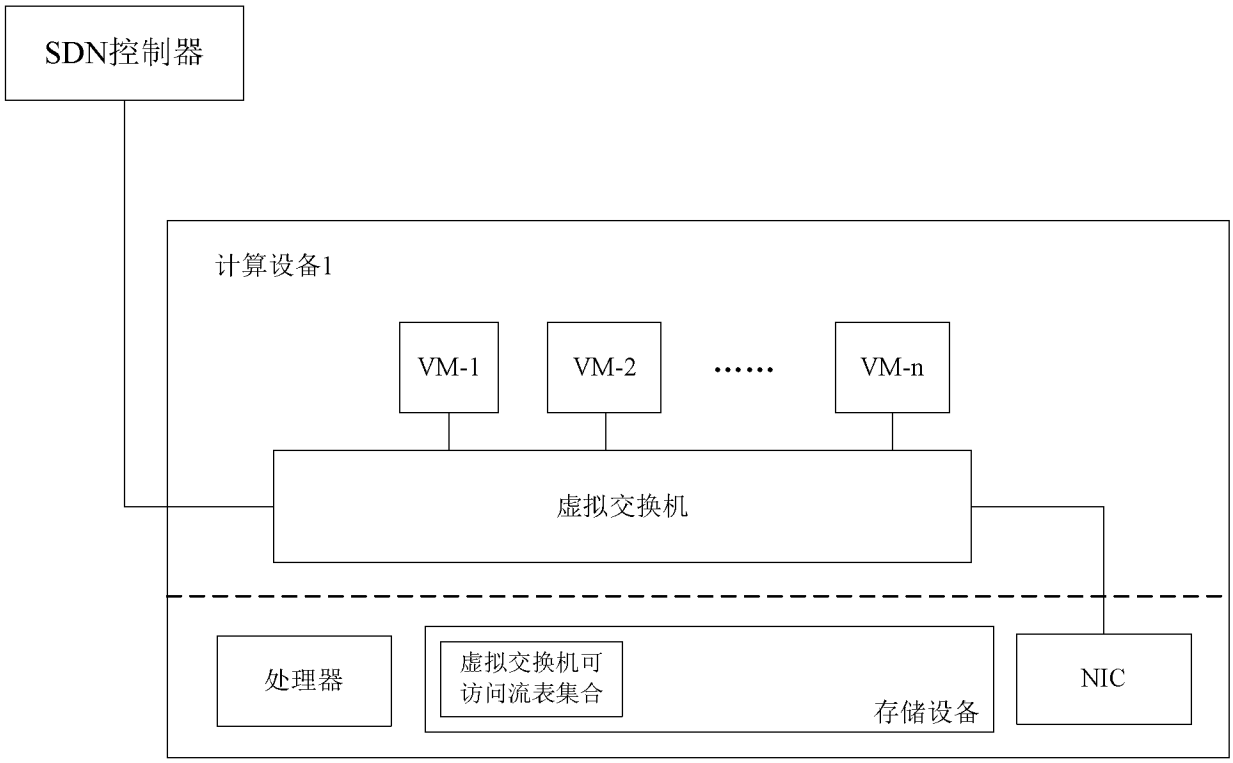
SDN控制器
计算设备1
VM-1
VM-2
……
VM-n
虚拟交换机
处理器
虚拟交换机可访问流表集合
存储设备
NIC

图 2b

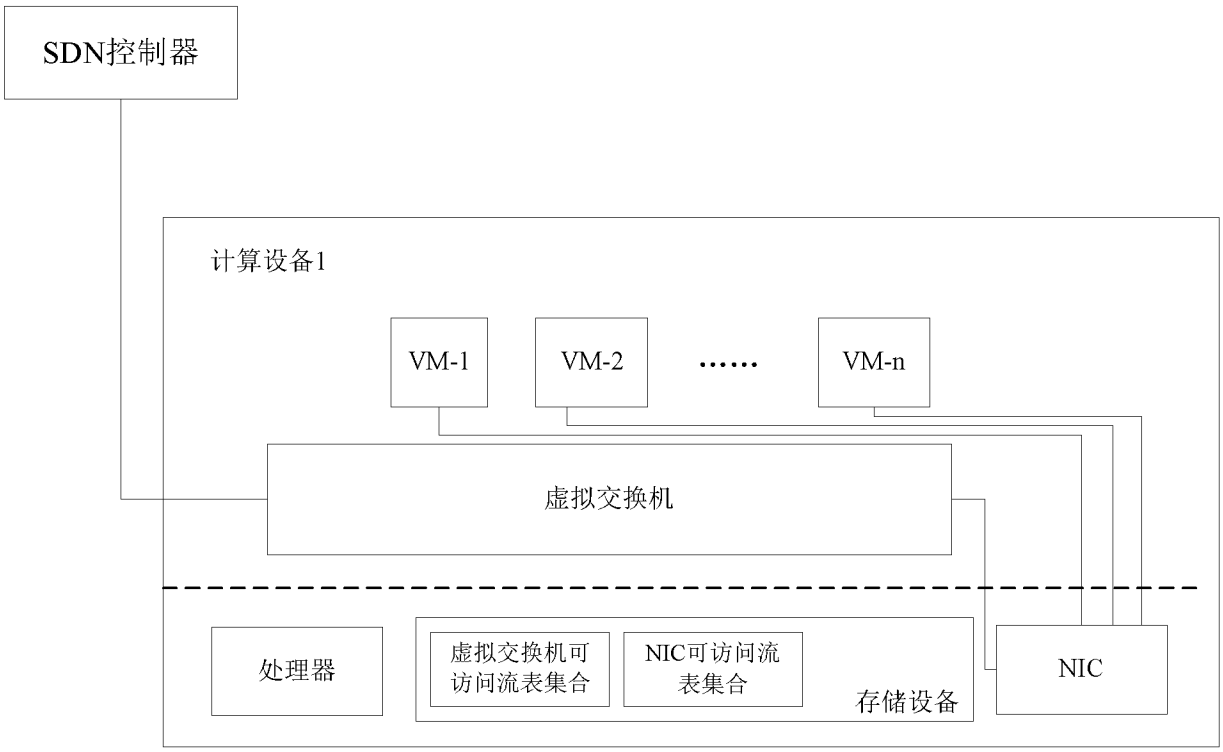
SDN控制器
计算设备1
VM-1
VM-2
……
VM-n
虚拟交换机
处理器
虚拟交换机可访问流表集合
NIC可访问流表集合
存储设备
NIC

图 2c

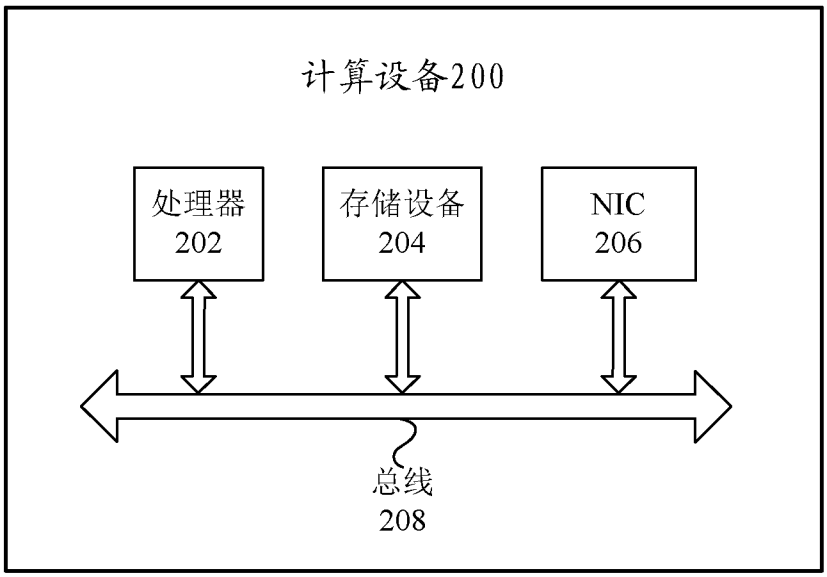

图 3

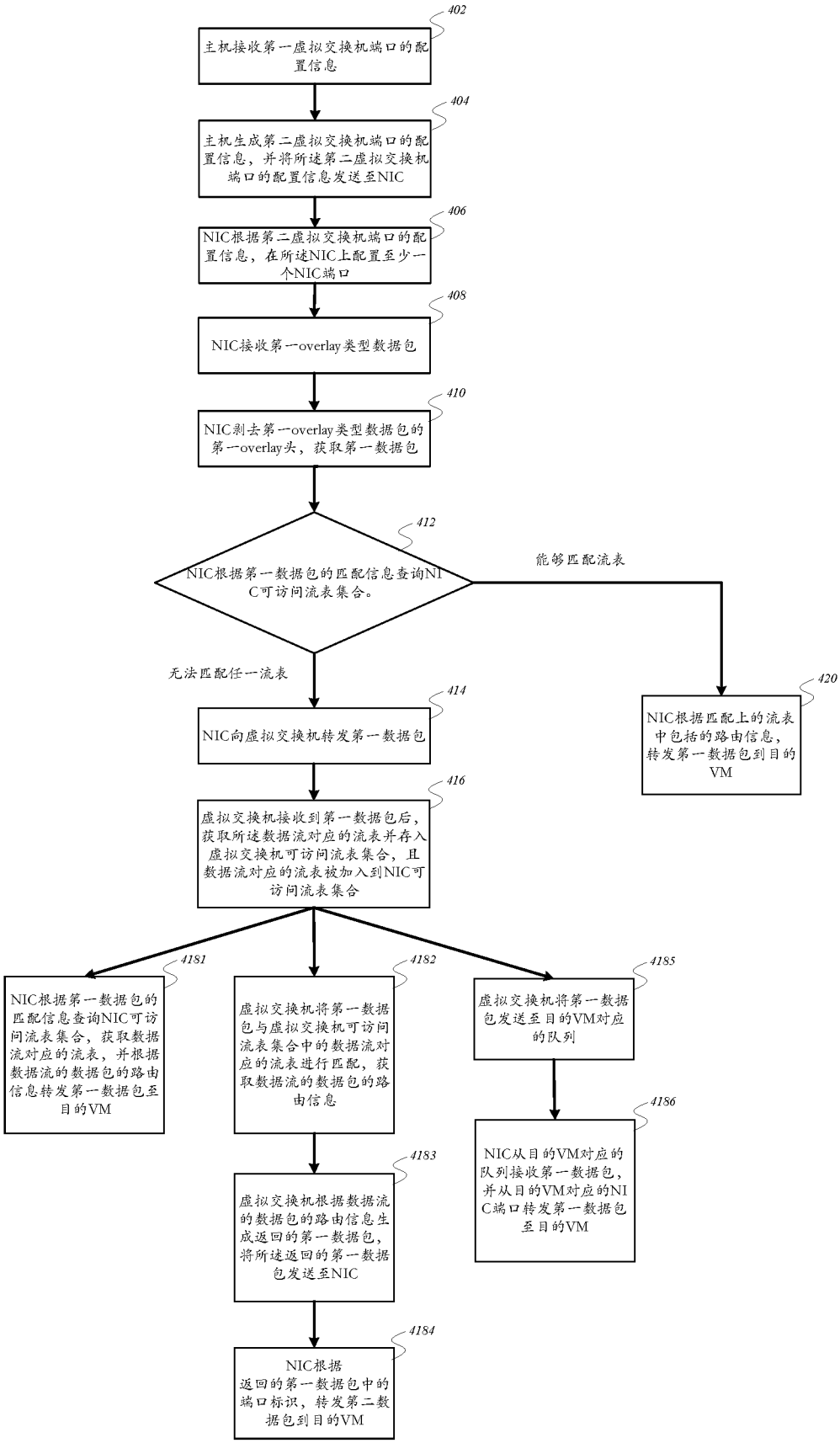
402
主机接收第一虚拟交换机端口的配置信息
404
主机生成第二虚拟交换机端口的配置信息，并将所述第二虚拟交换机端口的配置信息发送至NIC
406
NIC根据第二虚拟交换机端口的配置信息，在所述NIC上配置至少一个NIC端口
408
NIC接收第一overlay类型数据包
410
NIC剥去第一overlay类型数据包的第一overlay头，获取第一数据包
412
NIC根据第一数据包的匹配信息查询NIC可访问流表集合。
能够匹配流表
无法匹配任一流表
420
NIC根据匹配上的流表中包括的路由信息，转发第一数据包到目的VM
414
NIC向虚拟交换机转发第一数据包
416
虚拟交换机接收到第一数据包后，获取所述数据流对应的流表并存入虚拟交换机可访问流表集合，且数据流对应的流表被加入到NIC可访问流表集合
4181
NIC根据第一数据包的匹配信息查询NIC可访问流表集合，获取数据流对应的流表，并根据数据流的数据包的路由信息转发第一数据包至目的VM
4182
虚拟交换机将第一数据包与虚拟交换机可访问流表集合中的数据流对应的流表进行匹配，获取数据流的数据包的路由信息
4185
虚拟交换机将第一数据包发送至目的VM对应的队列
4186
NIC从目的VM对应的队列接收第一数据包，并从目的VM对应的NIC端口转发第一数据包至目的VM
4183
虚拟交换机根据数据流的数据包的路由信息生成返回的第一数据包，将所述返回的第一数据包发送至NIC
4184
NIC根据返回的第一数据包中的端口标识，转发第二数据包到目的VM

图 4

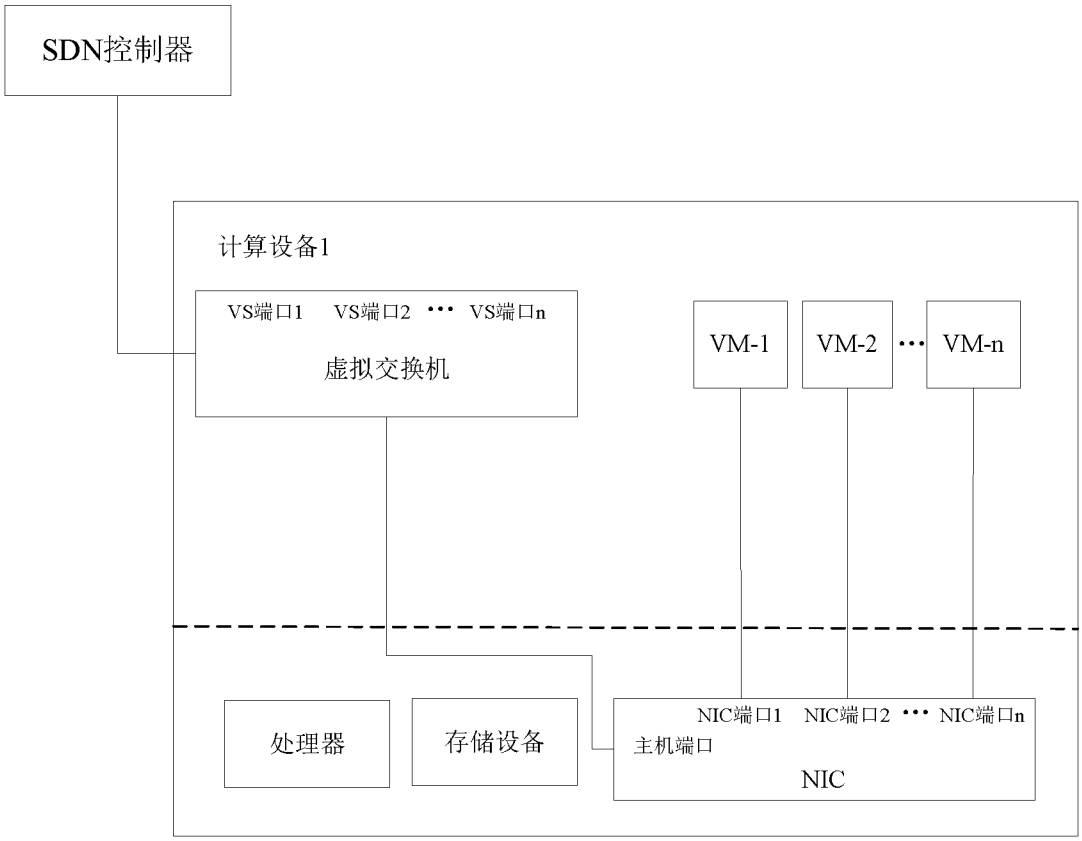
SDN控制器
计算设备1
VS端口1 VS端口2 ••• VS端口n
虚拟交换机
VM-1
VM-2
•••
VM-n
处理器
存储设备
NIC端口1 NIC端口2 ••• NIC端口n
主机端口
NIC

图 5a

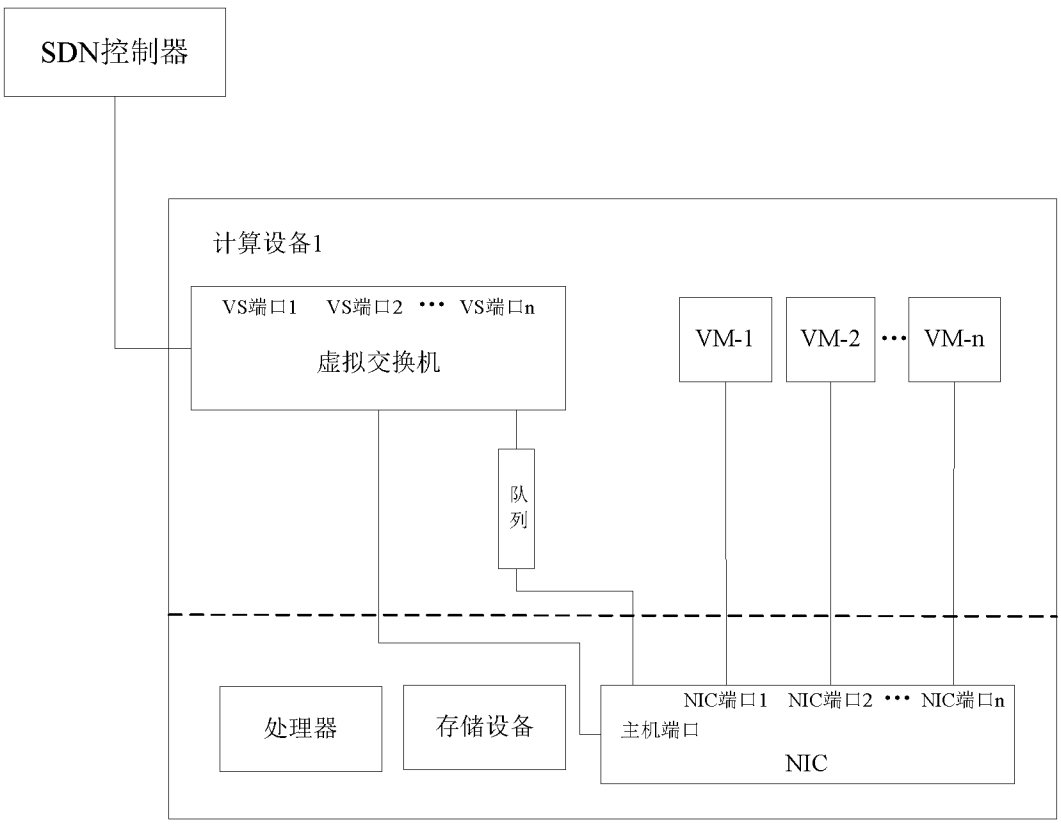
SDN控制器
计算设备1
VS端口1 VS端口2 ••• VS端口n
虚拟交换机
VM-1
VM-2
•••
VM-n
队列
处理器
存储设备
NIC端口1 NIC端口2 ••• NIC端口n
主机端口
NIC

图 5b

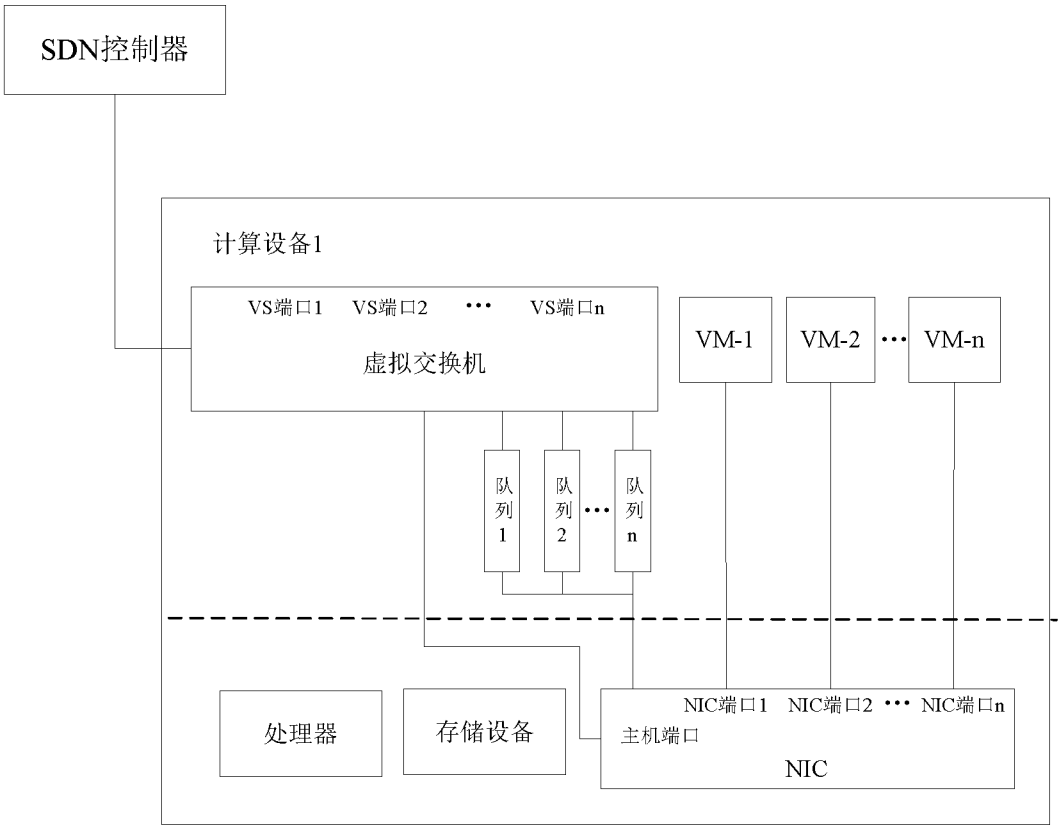
SDN控制器
计算设备1
VS端口1
VS端口2
VS端口n
虚拟交换机
VM-1
VM-2
VM-n
队列1
队列2
队列n
处理器
存储设备
NIC端口1
NIC端口2
NIC端口n
主机端口
NIC

图 5c

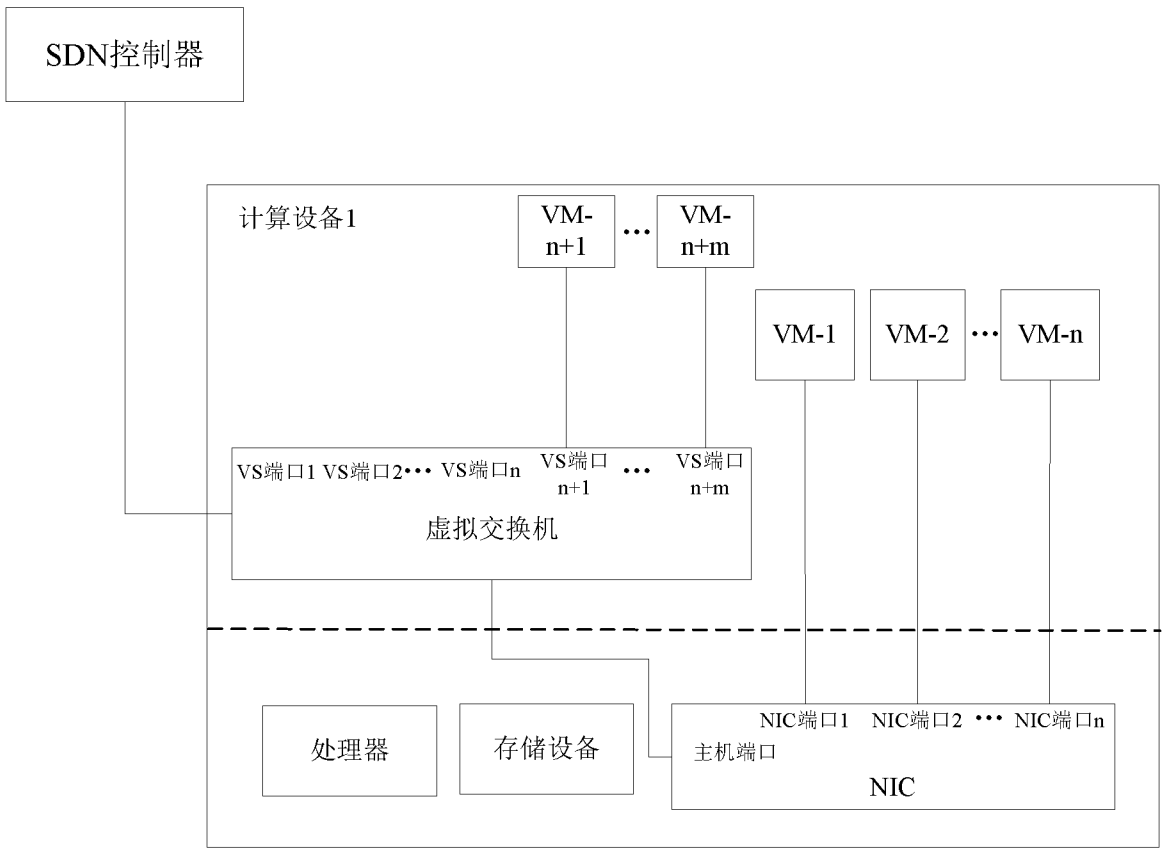
SDN控制器
计算设备1
VM-n+1
VM-n+m
VM-1
VM-2
VM-n
VS端口1
VS端口2
VS端口n
VS端口n+1
VS端口n+m
虚拟交换机
处理器
存储设备
NIC端口1
NIC端口2
NIC端口n
主机端口
NIC

图 5d

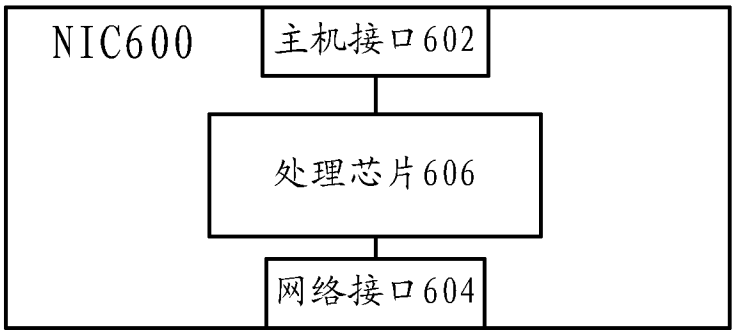
NIC600
主机接口602
处理芯片606
网络接口604

图 6a

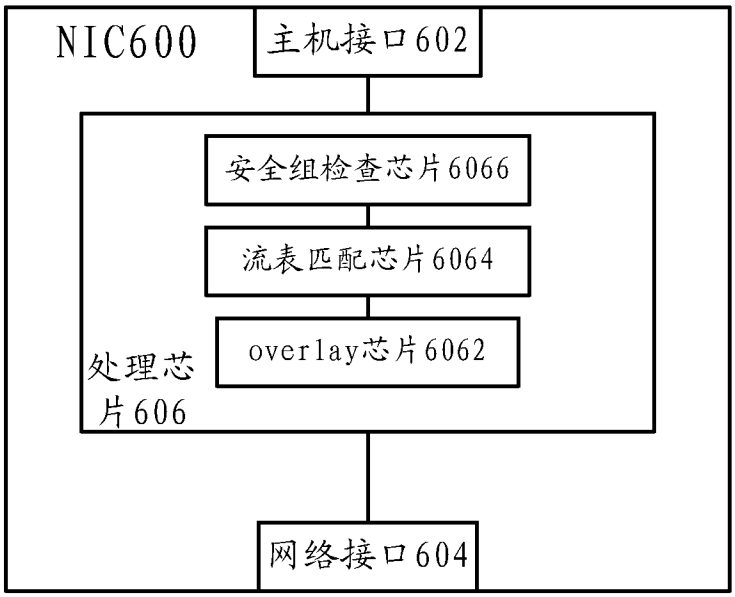
NIC600
主机接口602
安全组检查芯片6066
流表匹配芯片6064
overlay芯片6062
处理芯片606
网络接口604

图 6b

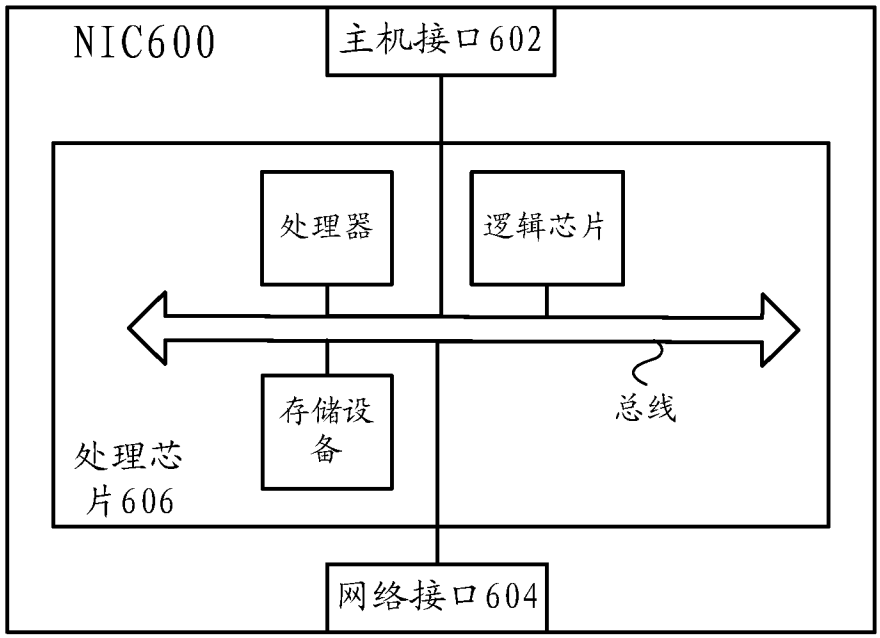

图 6c

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/093098

A. CLASSIFICATION OF SUBJECT MATTER

G06F 13/14 (2006.01) i; H04L 12/931 (2013.01) 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; H04L

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)

WPI; EPODOC; CNKI; IEEE; CNPAT: 网络接口，云，虚拟机，流表，匹配，路由, NIC, Openflow, flow table, match, rout+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104205080 A (MICROSOFT CORP.), 10 December 2014 (10.12.2014), description, paragraphs [0023]-[0045]	1-35
A	CN 103856573 A (HUAWEI TECHNOLOGIES CO., LTD.), 11 June 2014 (11.06.2014), entire document	1-35
A	CN 103346981 A (HUAWEI TECHNOLOGIES CO., LTD.), 09 October 2013 (09.10.2013), entire document	1-35
A	US 2010135182 A1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE), 03 June 2010 (03.06.2010), entire document	1-35

☐ 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

"E" earlier application or patent but published on or after the international filing date

"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)

"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

"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

"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

"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

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 March 2017	29 March 2017
Name and mailing address of the ISA/CN: 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 QIN, Xiaofang Telephone No.: (86-10) 62413241

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/093098

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104205080 A	10 December 2014	KR 20140143155 A	15 December 2014
		JP 2015515798 A	28 May 2015
		US 2013254766 A1	26 September 2013
		WO 2013142041 A1	26 September 2013
		EP 2828760 A1	28 January 2015
CN 103856573 A	11 June 2014	None	
CN 103346981 A	09 October 2013	WO 2014206105 A1	31 December 2014
		US 2015026681 A1	22 January 2015
		EP 2996294 A1	16 March 2016
US 2010135182 A1	03 June 2010	KR 20100062851 A	10 June 2010

国际检索报告

国际申请号
PCT/CN2016/093098

A. 主题的分类

G06F 13/14(2006.01)i; H04L 12/931(2013.01)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; H04L

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）)

WPI;EPODOC;CNKI;IEEE;CNPAT:网络接口，云，虚拟机，流表，匹配，路由，NIC，Openflow，flow table，match，rout+

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
X	CN 104205080 A (微软公司) 2014年 12月 10日 (2014-12-10) 说明书第[0023]段-第[0045]段	1-35
A	CN 103856573 A (华为技术有限公司) 2014年 6月 11日 (2014-06-11) 全文	1-35
A	CN 103346981 A (华为技术有限公司) 2013年 10月 9日 (2013-10-09) 全文	1-35
A	US 2010135182 A1 (ELECTRONICS AND TELECOMMUNICATIONS RESEARCH INSTITUTE) 2010年 6月 3日 (2010-06-03) 全文	1-35

☐ 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	国际检索报告邮寄日期
2017年 3月 7日	2017年 3月 29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 传真号 (86-10)62019451	受权官员 秦晓芳 电话号码 (86-10)62413241

表 PCT/ISA/210 (第2页) (2009年7月)

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2016/093098

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	104205080	A	2014年 12月 10日	KR	20140143155	A	2014年 12月 15日
				JP	2015515798	A	2015年 5月 28日
				US	2013254766	A1	2013年 9月 26日
				WO	2013142041	A1	2013年 9月 26日
				EP	2828760	A1	2015年 1月 28日
CN	103856573	A	2014年 6月 11日		无		
CN	103346981	A	2013年 10月 9日	WO	2014206105	A1	2014年 12月 31日
				US	2015026681	A1	2015年 1月 22日
				EP	2996294	A1	2016年 3月 16日
US	2010135182	A1	2010年 6月 3日	KR	20100062851	A	2010年 6月 10日