

# 云计算系统中报文处理的方法、主机和系统

## 技术领域

本发明涉及 IT 技术领域，尤其涉及云计算系统中报文处理的方法、主机和系统。

5

## 背景技术

虚拟化技术是云计算领域的关键技术之一，可将主机的物理资源抽象成共享资源池，供主机上运行的若干个虚拟机（Virtual Machine，VM）使用。主机上运行的若干个虚拟机可以共享使用主机的网卡，与主机外部网络进行通信。

10

在现有技术中，可以通过网卡虚拟化的方式将网卡分配给虚拟机使用，网卡虚拟化可以采用单根输入输出虚拟化（Single-Root I/O Virtualization，SR-IOV）或或多根输入输出虚拟化（Multi Root Input/Output Virtualization，MR-IOV），上述网卡虚拟化又称为网卡直通。以 SR-IOV 直通为例，当网卡支持 SR-IOV 时，使用 SR-IOV 技术可以将主机上网卡共享给主机上运行的若干个虚拟机。

15

在主机上使用支持 SR-IOV 能力的网卡时，网卡的网口会虚拟出至少一个物理功能（Physical Function，PF）以及多个虚拟功能（Virtual Function，VF），主机上的虚拟机连接到至少一个 VF。网卡内部包含具备交换机功能的交换设备。该交换设备按照介质访问控制（MediaAccessControl，MAC）表转发数据包，负责着 PF、VF 和物理网口之间的数据包转发，但由于该交换设备的处理和存储能力有限，因此无法为用户提供丰富的网络功能，例如安全组、服务质量（Quality of Service，QoS）、二层隧道封装以及分布式路由等等。

20

## 发明内容

本文描述了一种云计算系统中报文处理的方法、主机和系统，以解决网卡直通场景下无法提供丰富的网络功能的问题。

25

第一方面，本发明实施例提供了一种主机，至少一个网卡连接到该主机，

所述主机上运行有虚拟机监视器 VMM，VMM 上运行有第一虚拟机，所述 VMM 包含虚拟网桥以及虚拟网络功能模块，所述虚拟网络功能模块可以提供网络功能处理的能力。所述网卡包含交换设备以及至少两个网口，其中，第一网口支持网卡虚拟化能力，所述第一网口对应于至少一个物理功能 PF 和多个虚拟功能 VF，  
5 所述第一网口的至少一个 VF 直通给所述第一虚拟机，

第一虚拟机将数据包从与自身相连的 VF 发出，所述数据包的目的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址，由第一网口的交换设备根据数据包的目的 MAC 地址对数据包进行转发，将数据包发送给 VMM 上的虚拟网桥，通过该虚拟网桥将数据包传递到虚拟网络功能模块，由虚拟网络功能模块为该数据包  
10 提供丰富的网络功能。在网卡直通的场景下，虚拟机将数据包从直通的 VF 发出后，通过上述方式可以将数据包从网卡引导到 VMM 中的虚拟网络功能模块，利用该软件模块为用户提供丰富的网络功能。

在一种可能的设计中，所述主机上可以运行有多个虚拟机，每个虚拟机与所述第一网口的至少一个 VF 相连，所述第一网口的多个 VF 具有不同的虚拟局域网 VLAN 标识。通过为每个 VF 设置与其他 VF 不同的 VLAN 标识，保证带有 VLAN  
15 标识的广播报文只在 VLAN 标识对应的虚拟局域网传播，从而不会被发送给具有不同 VLAN 标识的其他 VF。

在一种可能的设计中，第一虚拟机通过 APR 流程获取待发送的数据包的目的 MAC 地址。具体的：

20 第一虚拟机通过与自身相连的 VF 发出 ARP 请求报文，所述 ARP 请求报文携带发出 ARP 请求报文的所述 VF 的 VLAN 标识，

第一网口的交换设备接收到所述 ARP 请求报文后，广播所述 ARP 请求报文，当 VMM 中的虚拟网桥接收到所述交换设备广播的所述 ARP 请求报文后，向 SDN 控制器转发所述 ARP 请求报文。SDN 控制器记录有云计算系统的网络拓扑信息，  
25 当 SDN 控制器收到 ARP 请求报文时，构建流表项，并向所述虚拟网桥返回所述流表项，所述流表项包括 ARP 代答设置，所述 ARP 代答设置用于指示所述虚拟网桥构造 ARP 响应报文，所述 ARP 响应报文携带所述虚拟网桥上行链路口的 MAC

地址作为所述待发送的数据包的目的 MAC 地址；

所述虚拟网桥根据所述流表项向所述第一虚拟机返回 ARP 响应报文，ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址。

- 5        本发明实施例定义了新的 ARP 流程，将待发送的数据包的目的 MAC 地址设定为虚拟网桥上行链路口的 MAC 地址，从而使得第一虚拟机发出的数据包经过二层转发到虚拟网桥，从而实现数据包从网卡重定向回 VMM 的效果。

在一种可能的设计中，所述流表项还包括 MAC 地址设置，所述 MAC 地址设置用于指示所述虚拟网桥在接收到第一虚拟机发送给第二虚拟机的所述数据包  
10    时，将所述数据包的目的 MAC 地址修改为第二虚拟机的 MAC 地址。由于第一虚拟机发出的数据包的目的 MAC 地址为虚拟网桥上行链路口的 MAC 地址，当虚拟网桥收到该数据包时，需要将数据包的目的 MAC 地址替换为第二虚拟机的 MAC 地址，从而使得虚拟网桥发出的数据包可以根据第二虚拟机真实的地址路由到第二虚拟机。

- 15        在一种可能的设计中，所述虚拟网桥上行链路口为所述第一网口的 PF，所述虚拟网桥通过所述 PF 与所述网卡相连；或者，所述虚拟网桥上行链路口为所述第一网口的选定的 VF，所述虚拟网桥通过所述选定的 VF 与所述网卡相连，其中，与所述虚拟网桥相连的所述选定的 VF 被设置为混杂模式。需要说明的是，所述虚拟网桥通过 PF 或 VF 与所述网卡相连指虚拟网桥与所述网卡存在逻辑相  
20    连，即所述虚拟网桥与所述网卡之间通过 PF 或 VF 传输数据包。

在一种可能的设计中，数据包的源虚拟机（第一虚拟机）和目的虚拟机（第二虚拟机）在同一主机上。此时，当虚拟网络功能模块处理完所述数据包后，将所述处理后的数据包返回给所述虚拟网桥，所述虚拟网桥接收处理后的数据包，根据所述数据包的目的 MAC 地址将所述处理后的数据包通过所述交换设备  
25    发送给所述第二虚拟机。

所述第二虚拟机与所述第一虚拟机也可以运行在不同主机上，此时，所述虚拟网络功能模块建立与所述第二虚拟机所在的主机上的另一虚拟网络功能模

块之间的隧道，将所述处理后的数据包发送到所述网卡的第二网口，然后通过所述隧道将所述处理后的数据包发送到所述第二虚拟机所在的主机，以使得所述另一虚拟网络功能模块将所述处理后的数据包发送到所述第二虚拟机。

第二方面，本发明实施例还提供了一种云计算系统中报文处理的方法，所述云计算系统中的至少一个主机包括虚拟机监视器 VMM，以及至少一个网卡，所述主机上运行有第一虚拟机，所述 VMM 包含虚拟网桥以及虚拟网络功能模块，所述网卡包含交换设备以及至少两个网口，其中，第一网口支持网卡虚拟化能力，所述第一网口对应于至少一个 PF 和多个 VF，所述第一虚拟机连接到所述第一网口的至少一个 VF，所述方法包括：

10 所述第一虚拟机确定待发送的数据包的目的 MAC 地址，所述待发送的数据包的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址；

所述第一虚拟机通过与自身相连的所述 VF 向第二虚拟机发送所述数据包，所述数据包携带的目的 IP 地址为所述第二虚拟机的 IP 地址，携带的目的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址；

15 所述交换设备接收所述数据包，根据所述数据包的目的 MAC 地址，将所述数据包通过所述虚拟网桥的上行链路口发送到所述虚拟网桥；

所述虚拟网桥接收所述数据包，将所述数据包的目的 MAC 地址修改为所述第二虚拟机的 MAC 地址，将修改后的数据包发送给虚拟网络功能模块；

20 所述虚拟网络功能模块对所述修改后的数据包进行网络功能处理后，发送处理后的数据包，所述处理后的数据包的目的 MAC 地址为所述第二虚拟机的 MAC 地址。

在一种可能的设计中，所述第一网口的多个 VF 具有不同的虚拟局域网 VLAN 标识。

25 在另一种可能的设计中，第一虚拟机通过 ARP 流程从 SDN 控制器获取待发送数据包的目的 MAC 地址，所述目的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址。具体获取的方式与第一方面的描述相同。

第三方面，本发明实施例提供了一种云计算系统，所述云计算系统包括第

一方面所述的主机以及 SDN 控制器，所述 SDN 控制器，用于接收所述虚拟网桥转发的来自所述第一虚拟机的 ARP 请求报文，向所述虚拟网桥返回流表项，所述流表项包括 ARP 代答设置和 MAC 地址设置，所述 ARP 代答设置用于指示所述虚拟网桥针对所述 ARP 请求构造 ARP 响应报文，所述 ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为待发送的数据包的目的 MAC 地址。

第四方面，与前述装置、方法以及系统相对应，本发明实施例提供了一种主机，该主机具有实现第一方面中定义的主机的功能。所述功能可以通过硬件实现，也可以通过硬件执行相应的软件实现。所述硬件或软件包括一个或多个与上述功能相对应的模块。具体的，所述主机包括第一处理器、第一存储器以及至少一个网卡，所述网卡包含第二处理器、第二存储器以及至少两个网口，其中，第一网口支持网卡虚拟化能力，所述第一网口对应于至少一个 PF 和多个 VF，所述第一虚拟机连接到所述第一网口的至少一个 VF，所述第一存储器和所述第二存储器中存储有指令，

所述第一处理器用于执行所述第一存储器中的指令以实施第一方面中第一虚拟机、虚拟网桥以及虚拟网络功能模块的功能；

所述第二处理器用于执行所述第二存储器中的指令以实施第一方面中交换设备的功能。

在一个可能的设计中，所述主机和 SDN 控制器由通用或专用的服务器实现，服务器的结构包括处理器、存储器、系统总线以及输入输出接口，所述处理器被配置为支持所述系统中主机的相应的功能。所述输入输出接口用于与所述云计算系统中的其他部件进行通信，处理器执行所述存储器中存储的指令。

第五方面，本发明实施例提供了一种计算机存储介质，用于储存为上述主机所用的计算机软件指令，其包含用于执行上述方面所设计的程序。

第六方面，本发明实施例提供了另一种计算机存储介质，用于储存为上述网卡中的交换设备所用的计算机软件指令，其包含用于执行上述方面所设计的程序。

第七方面，本发明实施例提供了一种计算机程序，当物理服务器运行该计

算机程序时，该物理服务器执行主机中虚拟机、虚拟网桥以及虚拟网络功能模块的功能。

第八方面，本发明实施例提供了一种计算机程序，当网卡中的处理器或可编程逻辑电路运行该计算机程序时，网卡执行前述方面中交换设备的功能。

5

## 附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍。显而易见地，下面附图中反映的仅仅是本发明的一部分实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得本发明的其他实施方式。而所有这些实施例或实施方式都在本发明的保护范围之内。

10

图 1A 是本发明实施例提供的主机上的虚拟化结构示意图；

图 1B 是本发明实施例提供的一种网卡直通场景下主机上的虚拟化结构示意图；

15

图 2 是本发明实施提供的另一种主机上的虚拟化架构示意图；

图 3 是本发明实施例提供的 ARP 流程示意图；

图 4 是本发明实施例提供的 VM1 向 VM2 发送数据包的流程示意图；

图 5 是本发明实施例提供的一种计算机设备硬件结构示意图；

图 6 是本发明实施例提供的一种云计算系统结构示意图。

20

## 具体实施方式

本发明实施例描述的网络架构以及业务场景是为了更加清楚的说明本发明实施例的技术方案，并不构成对于本发明实施例提供的技术方案的限定，本领域普通技术人员可知，随着网络架构的演变和新业务场景的出现，本发明实施例提供的技术方案对于类似的技术问题，同样适用。

25

如图 1A 所示，为本发明实施例提供的主机上的虚拟化结构示意图。主机为一台物理服务器，该物理服务器的底层为硬件层，硬件层主要包括中央处理

器 (CPU, Central Processing Unit)、内存、硬盘以及网卡等硬件资源。服务器虚拟化是在物理服务器上借助虚拟化软件 (如 VMWare ESX、Citrix XEN) 实现多个虚拟机 (Virtual Machine, VM) 的虚拟化运行环境。安装在服务器上实现虚拟化环境的软件层被称为虚拟机监视器 (Virtual Machine Monitor, VMM)。

- 5 运行在硬件层之上的 VMM 承担对硬件层中的硬件资源进行调度、分配和管理工作。VMM 之上运行多个虚拟机 VM, VMM 为每个虚拟机提供虚拟化的 CPU、内存、存储、IO 设备 (如网卡) 以及以太网交换机等硬件环境, 保证多个虚拟机相互隔离运行。

- 在虚拟化运行环境中, VMM 为每个虚拟机创建虚拟网卡 (Virtual Network Interface Card, vNIC), 虚拟交换机 VSwitch 提供了虚拟机之间, 以及虚拟机与外部网络之间的通讯能力。每个虚拟机的虚拟网卡对应到 VSwitch 的一个逻辑端口上, 主机的物理网卡对应于 VSwitch 与外部物理交换机相连的端口。当虚拟机发送或接收的数据包经过 VMM 时, VMM 中的虚拟网络功能模块对经过的数据进行网络功能处理, 由于虚拟网络功能模块为软件模块, 可以根据需要进行更新, 因此, VMM 中的虚拟网络功能模块可以为用户提供丰富的网络功能。
- 10
- 15

- 物理网卡的网口支持虚拟化能力可以通过单根输入输出虚拟化 (Single Root Input/Output Virtualization, SR-IOV) 或多根输入输出虚拟化 (Multi Root Input/Output Virtualization, MR-IOV) 来实现, 本发明实施例以 SR-IOV 技术为例进行说明。SR-IOV 技术是一种基于硬件的虚拟化解决方案, 可以在虚拟机之间高效共享 PCIe (Peripheral Component Interconnect Express, 快速外设组件互连) 设备, 并且由于 SR-IOV 技术是在硬件中实现的, 因此可以获得高效的 I/O 性能。
- 20

- 标准化组织 PCI-SIG (Peripheral Component Interconnect Special Interest Group, 外围设备互连专业组) 定义了 SR-IOV 规范, SR-IOV 规范可参阅 <http://www.pcisig.com>。
- 25

通过使用 SR-IOV 技术, 单个 I/O 资源可由主机上的多个虚拟机共享, 这样, 每个虚拟机都可访问同一个的硬件资源。因此, 启用了 SR-IOV 的 PCIe

设备（例如物理网卡的网口）可以显示为多个单独的设备，显示出的每个单独的设备都具有独立的 PCIe 配置空间。以支持 SR-IOV 的物理网卡为例，该物理网卡上包含若干个网口，可以分别针对各个网口启用/禁用 SR-IOV 能力，启用了 SR-IOV 能力的网口对应于至少一个物理功能 (Physical Function, PF) 以及  
5 多个虚拟功能 (Virtual Function, VF)，根据现有的 SR-IOV 规范，每个 PF 最多可有 64,000 个与其关联的 VF。创建 VF 后，可以直接将 VF 指定给主机上的虚拟机使用，从而使得多个虚拟机通过与其相连的至少一个 VF 共享 PCIe 设备。

如图 1B 所示，为本发明实施例提供的一种网卡直通场景下主机上的虚拟化  
10 结构示意图，与图 1A 不同的是，物理网卡支持单根 IO 虚拟化，物理网卡的网口对应于至少一个 PF 以及多个 VF，各 VF 能共享物理网卡的物理资源（例如网卡端口）。

当主机上的虚拟机通过直通 VF 发送数据包时，虚拟机发出的数据包将会通过虚拟机上安装的 VF 驱动直接传递给网卡的 VF，导致数据包无法经过 VMM 中的虚拟网络功能模块。由于网卡的交换设备中的 MAC 表容量有限，处理能力有限，且无法进行自由的功能扩展，因此网卡无法提供丰富的网络功能。在一种可能的场景中，交换设备为 virtual ethernet bridge & classifier。  
15

需要说明的是，本发明实施例针对网卡直通的场景，上述网卡直通的实现可以是 SR-IOV 或者 MR-IOV，本发明实施例对此并不进行限定。

20 如图 2 所示，为本发明实施例提供的另一种主机上的虚拟化架构示意图，该主机包含虚拟机监视器 VMM 以及至少一个物理网卡，所述主机上运行有至少一个虚拟机。所述至少一个物理网卡包括至少两个网口，第一网口和第二网口。其中，第一网口支持网卡虚拟化能力，第二网口连接到主机外部的物理交换机。第一网口虚拟出至少一个 PF 以及至少两个 VF（图 2 中以两个 VF 为例进行说明），  
25 第一虚拟机和第二虚拟机上安装有 VF 驱动，第一虚拟机和第二虚拟机分别连接到至少一个 VF，PF 连接到 VMM 上的虚拟网桥，该虚拟网桥与 VMM 上的虚拟网络功能模块相连，该虚拟网络功能模块通过第二网口与主机外部的物理交换机相

连。

在一种具体的场景中，VMM 上的虚拟网桥可以为 openvswitch bridge，将第一网口的 PF 作为该虚拟网桥的上行链路口。需要说明的是，也可以选择第一网口的一个 VF 作为虚拟网桥的上行链路口，即第一网口通过选择的 VF 与 VMM 上的虚拟网桥相连，此时，需要在 VMM 上加载 VF 驱动，同时将与虚拟网桥相连的 VF 设置为混杂模式，设置为混杂模式的 VF 类似于 PF，可以接收到目的地址不是自身 MAC 地址的数据包。混杂模式的具体说明可以参阅 [https://en.wikipedia.org/wiki/Promiscuous\\_mode](https://en.wikipedia.org/wiki/Promiscuous_mode)。本发明各实施例以第一网口的 PF 连接虚拟网桥为例进行说明，本领域技术人员可以理解的是，通过上述的设置方式，也可以使用混杂模式 (promiscuous mode) 的 VF 连接虚拟网桥。

在一种可能的实施方式中，第一网口和第二网口可以处在同一网卡上，也可以单独位于不同的网卡上，本发明实施例对此并不进行限定。

在本发明各实施例中，网卡虚拟化具体是指网卡硬件虚拟化。

第一虚拟机和第二虚拟机分别连接到第一网口的至少一个 VF，第一网口的每个 VF 设置有不同的虚拟局域网 (Virtual Local Area Network, VLAN) 标识，由于一个 VLAN 就是一个广播域，广播报文只能发送到到一个 VLAN 内的虚拟机。所以，通过为每个 VF 设置的不同的 VLAN 标识，使得每个虚拟机处在不同的广播域，虚拟机发出的广播报文只能被虚拟机自身和 PF 收到。在设置 VF 的 VLAN 标识时，需要保证主机范围内 VLAN 标识唯一即可，不同主机上的 VLAN 标识可以相同。

需要说明的是，图 2 所示的实施例中，第一虚拟机向第二虚拟机发送数据包是同一主机内部的两个虚拟机之间的通信。本领域技术人员可以理解的是，第二虚拟机与第一虚拟机也可以位于不同主机上，此时第一虚拟机和第二虚拟机为跨主机通信。图 2 中仅以两个 VF 以及两个 VLAN 子接口为例进行说明，在实际产品实现时，支持网卡虚拟化功能的网卡可以虚拟化为至少一个 PF 以及多个 VF，并不限定为两个 VF，相对应的，VLAN 子接口的数量也并限定为两个。

本发明实施例以 VM1 (第一虚拟机) 发送数据包到 VM2 (第二虚拟机) 为例

对数据包的处理流程进行说明，网卡中的交换设备需要将来自 VM1 的数据包发送到 VMM 上的虚拟网络功能模块进行处理。在发送数据包之前，第一虚拟机确定待发送的数据包的介质访问控制 MAC 地址为所述 VMM 中的虚拟网桥上行链路口的 MAC 地址。第一虚拟机通过与自身相连的所述 VF 向第二虚拟机发送所述数据包，所述数据包携带的目的 IP 地址为所述第二虚拟机的 IP 地址，携带的目的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址。第一网口的交换设备接收第一虚拟机通过 VF 发出的数据包，根据所述数据包的目的 MAC 地址对数据包进行二层转发，将所述数据包通过所述虚拟网桥的上行链路口发送到所述虚拟网桥。虚拟网桥接收所述数据包，将所述数据包的目的 MAC 地址修改为所述第二虚拟机的 MAC 地址，将修改后的数据包发送给虚拟网络功能模块，通过修改数据包的目的 MAC 地址，使得数据后续的传输过程可以使用修改后的 MAC 地址，从而保证数据包可以发送到第二虚拟机。虚拟网络功能模块在接收到修改后的数据包后，对所述修改后的数据包进行网络功能处理，为用户提供丰富的网络功能服务，然后根据所述第二虚拟机的 MAC 地址，将处理后的数据包发送到第二虚拟机。

本发明实施例中，虚拟机将数据包从与虚拟机相连的 VF 发出，由第一网口的交换设备进行转发，将数据包发送给 VMM 上的虚拟网桥，通过该虚拟网桥将数据包传递到虚拟网络功能模块，由虚拟网络功能模块为该数据包提供丰富的网络功能。在网卡直通的场景上，虚拟机将数据包从直通的 VF 发出后，通过上述方式可以将数据包发送到 VMM 中的虚拟网络功能模块，利用该软件模块为用户提供丰富的网络功能，实现对数据包进行虚拟网络功能处理。

进一步的，本发明实施例提供了一种地址解析协议（Address Resolution Protocol, ARP）流程，以使得第一虚拟机在发送数据包前可以获得虚拟网桥上行链路口的 MAC 地址，从而将数据包的目的 MAC 地址设定为虚拟网桥上行链路口的 MAC 地址。在本发明实施例提供的 ARP 流程中，由软件定义网络（Software Defined Network, SDN）控制器对 VM1 发起的 ARP 请求进行代答，在 ARP 响应中携带 VMM 上的虚拟网桥的上行链路口的 MAC 地址，通过将该 MAC 地址与数据

包的目的 IP 相对应，从而使得将 VM1 发出的数据包引导到与虚拟网桥相连的 PF 或者 VF。需要说明的是，SDN 控制器可以获得云计算网络的网络拓扑，即可实现上述 ARP 流程。

具体的，如图 3 所示，为本发明实施例提供的 ARP 流程示意图。

5        步骤 301：VM1 发出 ARP 请求报文，携带 VM2 的 IP 地址，用于获取 VM2 的 MAC 地址，该 ARP 请求报文从与 VM1 相连的 VF 发出，网卡上的交换设备接收到该 ARP 请求报文，该 ARP 请求报文携带发出所述 ARP 请求报文的 VF 的 VLAN 标识；

      步骤 302：交换设备广播该 ARP 请求报文，由于主机上的每个虚拟机连接的  
10    的 VF 都具备不同的 VLAN 标识，该 ARP 请求报文被通过 PF 发送给虚拟网桥；

      在一种可能的设计中，当虚拟网桥通过 VF 与网卡相连时，所述 ARP 请求报文通过 AF 发送到虚拟网桥。

      步骤 303：虚拟网桥将该 ARP 请求报文转发给 SDN 控制器；

      具体的，虚拟网桥可以通过 packet-in 消息将该 ARP 请求报文发送给 SDN  
15    控制器。

      步骤 304：SDN 控制器对虚拟网桥设置流表项，将流表项下发到虚拟网桥；  
      所述流表项包括：

      1) ARP 代答设置，该 ARP 代答设置表示：当虚拟网桥接收到 VM1 发出的用于获取 VM2 的 MAC 地址的 ARP 报文请求时，构造 ARP 响应报文，在 ARP 响应  
20    报文中携带虚拟网桥上行链路口的 MAC 地址作为 VM2 的 MAC 地址，该虚拟网桥上行链路口的 MAC 地址为与虚拟网桥相连的 PF 或 VF 的 MAC 地址；

      2) VLAN 标识处理设置：对于 VM1 发出的单播报文，去除单播报文携带的 VLAN 标识；

      3) MAC 地址设置：对于 VM1 发送给 VM2 的单播报文，将单播报文的目的  
25    MAC 地址修改为 VM2 的 MAC 地址，将修改后的单播报文通过与虚拟网络功能模块相连的端口发送到虚拟网络功能模块。

      需要说明的是，上述流表项的设置原理可以基于 Openflow 协议。

步骤 305: 虚拟网桥配置流表项, 根据 ARP 代答设置向 VM1 返回 ARP 响应报文, 该 ARP 响应报文携带虚拟网桥上行链路口的 MAC 地址, 将该 ARP 响应报文通过该上行链路口发送给交换设备;

步骤 306: 交换设备转发该 ARP 响应报文至 VM1;

5 步骤 307: VM1 接收该 ARP 响应报文, 将 VM2 的 MAC 地址设置为虚拟网桥上链路口的 MAC 地址。

本发明实施例定义了上述 ARP 过程, 无需通过自学习 MAC 转发表的形式实现 VM 流量的二层转发控制。通过上述 ARP 流程, 使得 VM1 发出的数据包的目的 MAC 为虚拟网桥上链路口的 MAC 地址, 从而保证了数据包通过该上行链路口  
10 发送到虚拟网桥, 进而由虚拟网桥将该数据包转发到虚拟网络功能模块进行处理。

如图 4 所示, 为本发明实施例提供的 VM1 向 VM2 发送数据包的流程示意图。在本发明实施例中, 主机 1 上的 VM1 向主机 2 上的 VM2 发送数据包, 主机 1 的 VMM 中包含第一虚拟网桥以及第一虚拟网络功能模块, 主机 1 的网卡包含至少  
15 两个网口以及第一交换设备, 主机 2 的 VMM 中包含第二虚拟网桥以及第二虚拟网络功能模块, 主机 2 的网卡包含至少两个网口以及第二交换设备。具体的, 主机 1 和主机 2 的结构和各组件的功能如图 2 所示, 此处不再赘述。VM1 向 VM2 发送数据包的流程包括:

步骤 401: VM1 通过与自身相连的 VF 发送目的为 VM2 的数据包, 数据包的目的 IP 地址为 VM2 的 IP 地址, 目的 MAC 地址为主机 1 上的第一虚拟网桥上  
20 链路口的 MAC 地址, 所述数据包携带所述 VF 的 VLAN 标识;

步骤 402: 网卡 1 的第一交换设备接收该数据包, 根据数据包的目的 MAC 地址将数据包通过该上行链路口发送给第一虚拟网桥;

步骤 403: 第一虚拟网桥根据预先设置的流表项, 对该数据包进行处理,  
25 去除该数据包 VLAN 标识, 将数据包的目的 MAC 地址修改为 VM2 的 MAC 地址;

步骤 404: 第一虚拟网桥将修改后的数据包发送给第一虚拟网络功能模块;

步骤 405: 第一虚拟网络功能模块对该数据包进行网络功能处理, 包括安

全组、QoS、二层隧道封装以及分布式路由等等；

需要说明的是，第一虚拟网络功能模块进行网络功能处理可以采用现有技术中的各种实现方式，本发明实施例对此并不进行限定。

步骤 406：第一虚拟网络功能模块将处理后的数据包通过第二网口发送到  
5 主机 1 外部的物理交换机，物理交换机经过网络将数据包路由到 VM2 所在的主机 2；

需要说明的是，数据包中携带有 VM2 的 IP 地址和 MAC 地址，采用现有的路由方式即可将数据包路由到 VM2 所在的主机 2，本发明实施例对数据包在网络中的传输过程并不进行限定，示例性的，VM1 所在的主机 1 和 VM2 所在的主机 2  
10 可以建立隧道，隧道的端点分别为主机 1 和主机 2 的虚拟网络功能模块，从而在两个主机之间为 VM1 和 VM2 建立数据传输所需的隧道。

步骤 407：VM2 所在主机 2 的第二虚拟网络功能模块对接收到的数据包进行网络功能处理，将处理后的数据包发送给第二虚拟网桥；

步骤 408：第二虚拟网桥根据 VM2 的 MAC 地址进行二层转发，通过交换设  
15 备将数据包发送给与 VM2 连接的 VF，从而将该数据包发送到第二虚拟机。

对于从主机外部发送到主机上的目的虚拟机的数据包，该数据包首先经过第二网口传递到虚拟网络功能模块，经过虚拟网络功能处理后再经过第一网口传递到数据包的目的虚拟机。

图 4 所示的实施例为虚拟机间跨主机的通信，可以理解的是，数据包的源  
20 虚拟机和目的虚拟机可以位于同一主机上，数据包的传输过程与图 4 的流程类似，不同之处在于，当数据包的源虚拟机和目的虚拟机位于同一主机上时，主机的虚拟网络功能模块完成网络功能处理后，将处理后的数据包发送给虚拟网桥，由虚拟网桥根据目的虚拟机的 MAC 地址将处理后的数据包发送给目的虚拟机。

25 主机、SDN 控制器可以采用通用的计算机设备，示例性的，如图 5 所示，为本发明实施例提供的计算机设备硬件结构示意图。计算机设备 500 包括至少一个处理器 501，通信总线 502，存储器 503 以及至少一个通信接口 504。

处理器 501 可以是一个通用中央处理器 (CPU), 微处理器, 特定应用集成电路 (application-specific integrated circuit, ASIC), 或一个或多个用于控制本发明方案程序执行的集成电路。

通信总线 502 可包括一通路, 在上述组件之间传送信息。所述通信接口 504, 5 使用任何收发器一类的装置, 用于与其他设备或通信网络通信, 如以太网, 无线接入网 (RAN), 无线局域网 (Wireless Local Area Networks, WLAN) 等。

存储器 503 可以是只读存储器 (read-only memory, ROM) 或可存储静态信息和指令的其他类型的静态存储设备, 随机存取存储器 (random access memory, RAM) 或者可存储信息和指令的其他类型的动态存储设备, 也可以是电 10 可擦可编程只读存储器 (Electrically Erasable Programmable Read-Only Memory, EEPROM)、只读光盘 (Compact Disc Read-Only Memory, CD-ROM) 或其他光盘存储、光碟存储 (包括压缩光碟、激光碟、光碟、数字通用光碟、蓝光光碟等)、磁盘存储介质或者其他磁存储设备、或者能够用于携带或存储具有指令或数据结构形式的期望的程序代码并能够由计算机存取的任何其他介质, 15 但不限于此。存储器可以是独立存在, 通过总线与处理器相连接。存储器也可以和处理器集成在一起。

其中, 所述存储器 503 用于存储执行本发明方案的应用程序代码, 并由处理器 501 来控制执行。所述处理器 501 用于执行所述存储器 503 中存储的应用程序代码。

20 在具体实现中, 作为一种实施例, 处理器 501 可以包括一个或多个 CPU, 例如图 5 中的 CPU0 和 CPU1。

在具体实现中, 作为一种实施例, 计算机设备 500 可以包括多个处理器, 例如图 5 中的处理器 501 和处理器 508。这些处理器中的每一个可以是一个单核 (single-CPU) 处理器, 也可以是一个多核 (multi-CPU) 处理器。这里的处 25 理器可以指一个或多个设备、电路、和/或用于处理数据 (例如计算机程序指令)

的处理核。

在具体实现中,作为一种实施例,计算机设备 500 还可以包括输出设备 505 和输入设备 506。输出设备 505 和处理器 501 通信,可以以多种方式来显示信息。例如,输出设备 505 可以是液晶显示器 (liquid crystal display, LCD),  
5 发光二极管 (light emitting diode, LED) 显示设备,阴极射线管 (cathode ray tube, CRT) 显示设备,或投影仪 (projector) 等。输入设备 506 和处理器 501 通信,可以以多种方式接受用户的输入。例如,输入设备 506 可以是鼠标、键盘、触摸屏设备或传感设备等。

上述的计算机设备 500 可以是一个通用计算机设备或者是一个专用计算机  
10 设备。在具体实现中,计算机设备 500 可以是台式机、便携式电脑、网络服务器、掌上电脑 (Personal Digital Assistant, PDA)、移动手机、平板电脑、无线终端设备、通信设备、嵌入式设备或有图 5 中类似结构的设备。本发明实施例不限定计算机设备 500 的类型。

SDN 控制器和图 2 中的主机可以为图 5 所示的设备,存储器中存储了一个  
15 或多个软件模块,用于实现主机和 SDN 控制器的各项功能。主机和 SDN 控制器可以通过处理器以及存储器中的程序代码来实现本发明实施例公开的报文处理的方法。

需要说明的是,图 5 所示的计算机设备仅仅是给出了云计算系统中各部分的可能的硬件实现方式,根据系统各部分功能的不同或者变化,可以对计算机  
20 设备的硬件组件进行增删,以使得与系统各部分的功能进行匹配。

进一步的,与图 5 所示的硬件装置类似,上述实施例中的网卡中包含处理器以及存储器,网卡中的处理器执行存储器中的指令以实现上述交换设备的功能。本发明实施例不再赘述。

更进一步的,如图 6 所示,为本发明实施例提供的一种云计算系统结构示意图,所述云计算系统包括至少一个主机 601 以及 SDN 控制器 602,所述主机  
25

中包含至少一个网卡，所述主机的结构如图 2 所示。

所述主机 601 包括虚拟机监视器 VMM，以及至少一个网卡，所述主机 601 上运行有第一虚拟机，所述 VMM 包含虚拟网桥以及虚拟网络功能模块，所述网卡包含交换设备以及至少两个网口，其中，第一网口支持单根输入输出虚拟化，  
5 所述第一网口对应于至少一个 PF 和多个 VF，所述第一虚拟机连接到所述第一网口的至少一个 VF，

所述 SDN 控制器 602，用于接收所述虚拟网桥转发的来自所述第一虚拟机的 ARP 请求报文，向所述虚拟网桥返回流表项，所述流表项包括 ARP 代答设置和 MAC 地址设置，所述 ARP 代答设置用于指示所述虚拟网桥为所述第一虚拟机  
10 构造 ARP 响应报文，所述 ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为待发送的数据包的目的 MAC 地址；

所述第一虚拟机，用于通过与自身相连的所述 VF 向第二虚拟机发送所述数据包，所述数据包携带的目的 IP 地址为所述第二虚拟机的 IP 地址，携带的目的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址；

15 所述交换设备，用于接收所述数据包，根据所述数据包的目的 MAC 地址，将所述数据包通过所述虚拟网桥的上行链路口发送到所述虚拟网桥；

所述虚拟网桥，用于接收所述数据包，将所述数据包的目的 MAC 地址修改为所述第二虚拟机的 MAC 地址，将修改后的数据包发送给虚拟网络功能模块；

所述虚拟网络功能模块，用于对所述修改后的数据包进行网络功能处理后，  
20 根据所述第二虚拟机的 MAC 地址，将处理后的数据包发送到第二虚拟机。

其中，所述第一网口的多个 VF 具有不同的虚拟局域网 VLAN 标识。

所述第一虚拟机，还用于通过所述与自身相连的所述 VF 发出所述 ARP 请求报文，所述 ARP 请求报文携带发出 ARP 请求报文的所述 VF 的 VLAN 标识；

所述交换设备，还用于广播所述 ARP 请求报文，所述 ARP 请求报文通过所述  
25 上行链路口发送到所述虚拟网桥；

所述虚拟网桥，还用于向所述 SDN 控制器转发所述 ARP 请求报文，接收所述 SDN 控制器返回的所述流表项，向所述第一虚拟机返回 ARP 响应报文，ARP

响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址。

所述流表项还包括 MAC 地址设置，所述 MAC 地址设置用于指示所述虚拟网桥在接收到第一虚拟机发送给第二虚拟机的单播报文时，将单播报文的目的  
5 MAC 地址修改为第二虚拟机的 MAC 地址。

进一步的流表项中还包括输出设置，所述输出设置用于指示将修改后的数据包通过与所述虚拟网络功能模块相连的端口发出。

所述虚拟网桥上行链路口为所述第一网口的 PF，所述虚拟网桥通过所述 PF 与所述网卡相连；或者，所述虚拟网桥上行链路口为所述第一网口的选定的 VF，  
10 所述虚拟网桥通过所述选定的 VF 与所述网卡相连，其中，与所述虚拟网桥相连的所述选定的 VF 被设置为混杂模式。

当所述第二虚拟机与所述第一虚拟机运行在同一主机上时，所述虚拟网桥，还用于接收处理后的数据包，根据所述数据包的目的 MAC 地址将所述处理后的数据包通过所述交换设备发送给所述第二虚拟机。

15 当所述第二虚拟机与所述第一虚拟机运行在不同主机上时，在第一虚拟机和第二虚拟机所在的两个主机的虚拟网络功能模块之间建立隧道，通过隧道将处理后的数据包从第二网口发出，传输到第二虚拟机所在的主机的虚拟网络功能模块。在第二虚拟机所在的虚拟网络功能模块对数据包完成网络功能处理后，将数据包通过第二虚拟机所在的主机的虚拟网桥发送到第二虚拟机。

20 本发明实施例还提供了一种计算机存储介质，用于储存为上述图 2-6 中的设备所用的计算机软件指令，其包含用于执行上述方法实施例所设计的程序。通过执行存储的程序，可以实现云计算系统中报文处理的方法。

综上所述，本发明实施例提供的云计算系统中报文处理的方法，主机和系统，可以将使用直通 VF 的虚拟机发出的报文引导到 VMM 上的虚拟网桥，进而由  
25 VMM 上的虚拟网络功能模块对数据包进行网络功能处理，从而为用户提供丰富的网络功能。

尽管在此结合各实施例对本发明进行了描述，然而，在实施所要求保护的

本发明实施例过程中，本领域技术人员通过查看所述附图、公开内容、以及所附权利要求书，可理解并实现所述公开实施例的其他变化。在权利要求中，“包括”（comprising）一词不排除其他组成部分或步骤，。单个处理器或其他单元可以实现权利要求中列举的若干项功能。相互不同的从属权利要求中记载了某些措施，但这并不表示这些措施不能组合起来产生良好的效果。

本领域技术人员应明白，本发明的实施例可提供为方法、装置（设备）、或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本发明实施例可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。计算机程序存储/分布在合适的介质中，与其它硬件一起提供或作为硬件的一部分，也可以采用其他分布形式，如通过 Internet 或其它有线或无线电信系统。

本发明实施例涉及到的计算机程序指令可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，通过执行计算机程序指令可以实现前述各实施例中各部件的功能

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

尽管结合具体特征及其实施例对本发明进行了描述，显而易见的，在不脱离本发明的精神和范围的情况下，可对其进行各种修改和组合。相应地，本说明书和附图仅仅是所附权利要求所界定的本发明的示例性说明，且视为已覆盖本发明范围内的任意和所有修改、变化、组合或等同物。显然，本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围。这样，倘若本发明的这些修改和变型属于本发明权利要求及其等同技术的范围之内，则本发明也意图包含这些改动和变型在内。

## 权 利 要 求

1. 一种主机，其特征在于，所述主机包括虚拟机监视器 VMM，以及至少一个网卡，所述主机上运行有第一虚拟机，所述 VMM 包含虚拟网桥以及虚拟网络功能模块，所述至少一个网卡包含交换设备以及至少两个网口，其中，第一网口支持网卡虚拟化能力，所述第一网口对应于至少一个物理功能 PF 和多个虚拟功能 VF，所述第一虚拟机连接到所述第一网口的至少一个 VF，

所述第一虚拟机，用于确定待发送的数据包的目的介质访问控制 MAC 地址，所述待发送的数据包的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址；

所述第一虚拟机，还用于通过与自身相连的所述 VF 向第二虚拟机发送所述数据包，所述数据包携带的目的 IP 地址为所述第二虚拟机的 IP 地址，携带的目的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址；

所述交换设备，用于接收所述数据包，根据所述数据包的目的 MAC 地址，将所述数据包通过所述虚拟网桥的上行链路口发送到所述虚拟网桥；

所述虚拟网桥，用于接收所述数据包，将所述数据包的目的 MAC 地址修改为所述第二虚拟机的 MAC 地址，将修改后的数据包发送给虚拟网络功能模块；

所述虚拟网络功能模块，用于对所述修改后的数据包进行网络功能处理后，发送处理后的数据包，所述处理后的数据包的目的 MAC 地址为所述第二虚拟机的 MAC 地址。

2. 如权利要求 1 所述的主机，其特征在于，所述第一网口的多个 VF 具有不同的虚拟局域网 VLAN 标识。

3. 如权利要求 2 所述的主机，其特征在于，

所述第一虚拟机，还用于通过所述与自身相连的所述 VF 发出地址解析协议 ARP 请求报文，所述 ARP 请求报文携带发出 ARP 请求报文的所述 VF 的 VLAN 标识；

所述交换设备，还用于广播所述 ARP 请求报文，所述 ARP 请求报文通过所述上行链路口发送到所述虚拟网桥；

所述虚拟网桥，还用于向软件定义网络 SDN 控制器转发所述 ARP 请求报文，

接收所述 SDN 控制器返回的流表项, 所述流表项包括 ARP 代答设置, 所述 ARP 代答设置用于指示所述虚拟网桥针对所述 ARP 请求报文构造 ARP 响应报文, 所述 ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址;

- 5       所述虚拟网桥, 还用于向所述第一虚拟机返回 ARP 响应报文, ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址。

4. 如权利要求 3 所述的主机, 其特征在于, 所述流表项还包括 MAC 地址设置, 所述 MAC 地址设置用于指示所述虚拟网桥在接收到第一虚拟机发送给第二  
10   虚拟机的所述数据包时, 将所述数据包的目的 MAC 地址修改为第二虚拟机的 MAC 地址。

5. 如权利要求 1-4 任一所述的主机, 其特征在于,

所述虚拟网桥上行链路口为所述第一网口的 PF, 所述虚拟网桥通过所述 PF 与  
所述网卡相连; 或者,

- 15       所述虚拟网桥上行链路口为所述第一网口的选定的 VF, 所述虚拟网桥通过所述选定的 VF 与  
所述网卡相连, 其中, 与所述虚拟网桥相连的所述选定的 VF 被设置为混杂模式。

6. 如权利要求 1-4 任一所述的主机, 其特征在于, 所述第二虚拟机与所述第一虚拟机运行在同一主机上,

- 20       所述虚拟网桥, 还用于接收处理后的数据包, 根据所述数据包的目的 MAC 地址将所述处理后的数据包通过所述交换设备发送给所述第二虚拟机。

7. 如权利要求 1-4 任一所述的主机, 其特征在于, 所述第二虚拟机与所述第一虚拟机运行在不同主机上,

- 所述虚拟网络功能模块, 具体用于建立与所述第二虚拟机所在的主机上的  
25   另一虚拟网络功能模块之间的隧道, 将所述处理后的数据包发送到所述网卡的第二网口, 通过所述隧道将所述处理后的数据包发送到所述第二虚拟机所在的主机, 以使得所述另一虚拟网络功能模块将所述处理后的数据包发送到所述第

二虚拟机。

8. 一种云计算系统中报文处理的方法，其特征在于，所述云计算系统中的至少一个主机包括虚拟机监视器 VMM，以及至少一个网卡，所述主机上运行有第一虚拟机，所述 VMM 包含虚拟网桥以及虚拟网络功能模块，所述网卡包含交换设备以及至少两个网口，其中，第一网口支持网卡虚拟化能力，所述第一网口对应于至少至少一个 PF 和多个 VF，所述第一虚拟机连接到所述第一网口的至少一个 VF，所述方法包括：

所述第一虚拟机确定待发送的数据包的目的 MAC 地址，所述待发送的数据包的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址；

10 所述第一虚拟机通过与自身相连的所述 VF 向第二虚拟机发送所述数据包，所述数据包携带的目的 IP 地址为所述第二虚拟机的 IP 地址，携带的目的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址；

所述交换设备接收所述数据包，根据所述数据包的目的 MAC 地址，将所述数据包通过所述虚拟网桥的上行链路口发送到所述虚拟网桥；

15 所述虚拟网桥接收所述数据包，将所述数据包的目的 MAC 地址修改为所述第二虚拟机的 MAC 地址，将修改后的数据包发送给虚拟网络功能模块；

所述虚拟网络功能模块对所述修改后的数据包进行网络功能处理后，发送处理后的数据包，所述处理后的数据包的目的 MAC 地址为所述第二虚拟机的 MAC 地址。

20 9. 如权利要求 8 所述的方法，其特征在于，所述第一网口的多个 VF 具有不同的虚拟局域网 VLAN 标识。

10. 如权利要求 9 所述的方法，其特征在于，所述第一虚拟机确定待发送的数据包的目的 MAC 地址包括：

25 所述第一虚拟机通过所述与自身相连的所述 VF 发出 ARP 请求报文，所述 ARP 请求报文携带发出 ARP 请求报文的所述 VF 的 VLAN 标识；

所述交换设备广播所述 ARP 请求报文，所述 ARP 请求报文通过所述上行链路口发送到所述虚拟网桥；

所述虚拟网桥向 SDN 控制器转发所述 ARP 请求报文，接收所述 SDN 控制器返回的流表项，所述流表项包括 ARP 代答设置，所述 ARP 代答设置用于指示所述虚拟网桥针对所述 ARP 请求报文构造 ARP 响应报文，所述 ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址；

5       所述虚拟网桥，还用于向所述第一虚拟机返回 ARP 响应报文，ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址。

11. 如权利要求 10 所述的方法，其特征在于，所述流表项还包括 MAC 地址设置所述 MAC 地址设置用于指示所述虚拟网桥在接收到第一虚拟机发送给第二  
10   虚拟机的所述数据包时，将所述数据包的目的 MAC 地址修改为第二虚拟机的 MAC 地址。

12. 如权利要求 8-11 任一所述的方法，其特征在于，

所述虚拟网桥上行链路口为所述第一网口的 PF，所述虚拟网桥通过所述 PF 与  
所述网卡相连；或者，

15       所述虚拟网桥上行链路口为所述第一网口的选定的 VF，所述虚拟网桥通过所述选定的 VF 与  
所述网卡相连，其中，与所述虚拟网桥相连的所述选定的 VF 被设置为混杂模式。

13. 如权利要求 8-11 任一所述的方法，其特征在于，所述第二虚拟机与所述第一虚拟机运行在同一主机上，所述方法还包括：

20       所述虚拟网桥接收处理后的数据包，根据所述数据包的目的 MAC 地址将所述处理后的数据包通过所述交换设备发送给所述第二虚拟机。

14. 如权利要求 8-11 任一所述的方法，其特征在于，所述第二虚拟机与所述第一虚拟机运行在不同主机上，所述方法还包括：

25       所述虚拟网络功能模块建立与所述第二虚拟机所在的主机上的另一虚拟网络功能模块之间的隧道，将所述处理后的数据包发送到所述网卡的第二网口，通过所述隧道将所述处理后的数据包发送到所述第二虚拟机所在的主机，以使得所述另一虚拟网络功能模块将所述处理后的数据包发送到所述第二虚拟机。

15. 一种云计算系统，其特征在于，包括软件定义网络 SDN 控制器，以及至少一个主机，所述主机包括虚拟机监视器 VMM，以及至少一个网卡，所述主机上运行有第一虚拟机，所述 VMM 包含虚拟网桥以及虚拟网络功能模块，所述网卡包含交换设备以及至少两个网口，其中，第一网口支持网卡虚拟化能力，所述第一网口对应于至少一个 PF 和多个 VF，所述第一虚拟机连接到所述第一网口的至少一个 VF，

所述 SDN 控制器，用于接收所述虚拟网桥转发的来自所述第一虚拟机的 ARP 请求报文，向所述虚拟网桥返回流表项，所述流表项包括 ARP 代答设置和 MAC 地址设置，所述 ARP 代答设置用于指示所述虚拟网桥为所述第一虚拟机构造 ARP 响应报文，所述 ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为待发送的数据包的目的 MAC 地址；

所述第一虚拟机，用于通过与自身相连的所述 VF 向第二虚拟机发送所述数据包，所述数据包携带的目的 IP 地址为所述第二虚拟机的 IP 地址，携带的目的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址；

15 所述交换设备，用于接收所述数据包，根据所述数据包的目的 MAC 地址，将所述数据包通过所述虚拟网桥的上行链路口发送到所述虚拟网桥；

所述虚拟网桥，用于接收所述数据包，将所述数据包的目的 MAC 地址修改为所述第二虚拟机的 MAC 地址，将修改后的数据包发送给虚拟网络功能模块；

20 所述虚拟网络功能模块，用于对所述修改后的数据包进行网络功能处理后，发送处理后的数据包，所述处理后的数据包的目的 MAC 地址为所述第二虚拟机的 MAC 地址。

16. 如权利要求 15 所述的系统，其特征在于，所述第一网口的多个 VF 具有不同的虚拟局域网 VLAN 标识。

17. 如权利要求 16 所述的系统，其特征在于，

25 所述第一虚拟机，还用于通过所述与自身相连的所述 VF 发出所述 ARP 请求报文，所述 ARP 请求报文携带发出 ARP 请求报文的所述 VF 的 VLAN 标识；

所述交换设备，还用于广播所述 ARP 请求报文，所述 ARP 请求报文通过所

述上行链路口发送到所述虚拟网桥；

所述虚拟网桥，还用于向所述 SDN 控制器转发所述 ARP 请求报文，接收所述 SDN 控制器返回的所述流表项，所述流表项包括 ARP 代答设置，所述 ARP 代答设置用于指示所述虚拟网桥针对所述 ARP 请求报文构造 ARP 响应报文，所述  
5 ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址；

所述虚拟网桥，还用于向所述第一虚拟机返回 ARP 响应报文，ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址。

10 18. 如权利要求 17 所述的系统，其特征在于，所述流表项还包括 MAC 地址设置，所述 MAC 地址设置用于指示所述虚拟网桥在接收到第一虚拟机发送给第二虚拟机的所述数据包时，将所述数据包的目的 MAC 地址修改为第二虚拟机的 MAC 地址。

19. 如权利要求 15-18 任一所述的系统，其特征在于，

15 所述虚拟网桥上行链路口为所述第一网口的 PF，所述虚拟网桥通过所述 PF 与所述网卡相连；或者，

所述虚拟网桥上行链路口为所述第一网口的选定的 VF，所述虚拟网桥通过所述选定的 VF 与所述网卡相连，其中，与所述虚拟网桥相连的所述选定的 VF 被设置为混杂模式。

20 20. 如权利要求 15-18 任一所述的系统，其特征在于，所述第二虚拟机与所述第一虚拟机运行在同一主机上，

所述虚拟网桥，还用于接收处理后的数据包，根据所述数据包的目的 MAC 地址将所述处理后的数据包通过所述交换设备发送给所述第二虚拟机。

21. 如权利要求 15-18 任一所述的系统，其特征在于，所述第二虚拟机与  
25 所述第一虚拟机运行在不同主机上，

所述虚拟网络功能模块，具体用于建立与所述第二虚拟机所在的主机上的另一虚拟网络功能模块之间的隧道，将所述处理后的数据包发送到所述网卡的

第二网口，通过所述隧道将所述处理后的数据包发送到所述第二虚拟机所在的主机，以使得所述另一虚拟网络功能模块将所述处理后的数据包发送到所述第二虚拟机。

22. 一种主机，其特征在于，包括第一处理器、第一存储器以及至少一个网  
5 卡，所述网卡包含第二处理器、第二存储器以及至少两个网口，其中，第一网口支持网卡虚拟化能力，所述第一网口对应于至少一个 PF 和多个 VF，所述第一存储器和所述第二存储器中存储有指令，所述第一处理器执行所述第一存储器中的第一指令以实现第一虚拟机的功能，所述第一虚拟机连接到所述第一网口的至少一个 VF，所述第一处理器执行所述第一存储器中的第二指令以实现虚拟网桥的功能，所述  
10 第一处理器执行所述第一存储器中的第三指令以实现虚拟网络功能模块的功能，所述第二处理器用于执行所述第二存储器中的指令以实现交换设备的功能，

所述第一处理器用于执行所述第一存储器中的第一指令以执行步骤：确定待发送的数据包的目的 MAC 地址，所述待发送的数据包的 MAC 地址为所述虚拟  
15 网桥上行链路口的 MAC 地址，通过与所述第一虚拟机相连的 VF 向第二虚拟机发送所述数据包，所述数据包携带的目的 IP 地址为所述第二虚拟机的 IP 地址，携带的目的 MAC 地址为所述虚拟网桥上行链路口的 MAC 地址；

所述第二处理器用于执行所述第二存储器中的指令以执行步骤：接收所述数据包，根据所述数据包的目的 MAC 地址，将所述数据包通过所述虚拟网桥的  
20 上行链路口发送到所述虚拟网桥；

所述第一处理器还用于执行所述第一存储器中的第二指令以执行步骤：接收所述数据包，将所述数据包的目的 MAC 地址修改为所述第二虚拟机的 MAC 地址；

所述第一处理器还用于执行所述第一存储器中的第三指令以执行步骤：对  
25 修改后的数据包进行网络功能处理后，发送处理后的数据包，所述处理后的数据包的目的 MAC 地址为所述第二虚拟机的 MAC 地址。

23. 如权利要求 22 所述的主机，其特征在于，述第一网口的多个 VF 具有不

同的虚拟局域网 VLAN 标识，

所述第一处理器还用于执行所述第一存储器中的所述第一指令以执行步骤：通过与所述第一虚拟机相连的所述 VF 发出 ARP 请求报文，所述 ARP 请求报文携带发出 ARP 请求报文的所述 VF 的 VLAN 标识；

- 5       所述第二处理器用于执行所述第二存储器中的指令以执行步骤：广播所述 ARP 请求报文，所述 ARP 请求报文通过所述上行链路口发送到所述虚拟网桥；

所述第一处理器还用于执行所述第一存储器中的第二指令以执行步骤：向 SDN 控制器转发所述 ARP 请求报文，接收所述 SDN 控制器返回的流表项，所述流表项包括 ARP 代答设置，所述 ARP 代答设置用于指示所述虚拟网桥针对所述  
10   ARP 请求报文构造 ARP 响应报文，所述 ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址；向所述第一虚拟机返回 ARP 响应报文，ARP 响应报文携带所述虚拟网桥上行链路口的 MAC 地址作为所述待发送的数据包的目的 MAC 地址。

24. 如权利要求 22 所述的主机，其特征在于，所述流表项还包括 MAC 地址  
15   设置所述 MAC 地址设置用于指示所述虚拟网桥在接收到第一虚拟机发送给第二虚拟机的所述数据包时，将所述数据包的目的 MAC 地址修改为第二虚拟机的 MAC 地址。